

**L'INDAGINE SUI CONSUMI:
ALCUNE VERIFICHE SU PROPOSTE
DI CAMPIONAMENTO ALTERNATIVE**

Daniela Cocchi* e Michele Lalla**

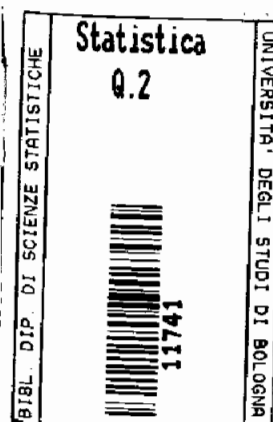
Rapporto di ricerca n. 12

CON PRI - La misura dei consumi privati

* Dipartimento di Scienze Statistiche "Paolo Fortunati"
Università di Bologna



Dipartimento di Scienze Statistiche "Paolo Fortunati"
dell'Università degli Studi di Bologna
Dicembre 1995



I lavori raccolti in questa collana hanno avuto origine nell'ambito del progetto di ricerca dell'Istat «Le statistiche dei consumi privati nel sistema statistico nazionale» e del progetto di ricerca MURST 40% «La misura dei consumi privati: uno studio sull'accuratezza, coerenza e qualità dei dati». Al progetto di ricerca hanno partecipato i ricercatori dell'Istat e dei seguenti Dipartimenti e Istituti universitari:

- Dipartimento di Scienze Statistiche, Bologna
- Dipartimento di Contabilità Nazionale, Roma
- Dipartimento Statistico, Firenze
- Istituto di Statistica e Matematica, Istituto Universitario Navale, Napoli
- Dipartimento di Scienze Statistiche, Perugia
- Istituto di Statistica, Messina.

Il lavoro è frutto della collaborazione dei due autori. Daniela Cocchi ha curato i capitoli 1 e 4, e i paragrafi 3.1, 3.4, e 5.2; Michele Lalla ha curato i capitoli 2, e 6, e i paragrafi 3.2, 3.3, 5.2 e 5.3.

Finito di stampare nel mese di Febbraio 1996
presso le Officine Grafiche TECNOPRINT S.N.C.
Via del Legatore 3, Bologna

INDICE

1. Introduzione	p. 5
2. Caratteristiche dell'indagine attuale sui consumi	p. 7
2.1. Alcune sintesi dell'indagine sui consumi del 1989	p. 8
2.1.1. Alcune analisi della varianza	p. 10
2.1.2. Un'analisi dei gruppi	p. 12
3. Proposte di modifiche al piano di campionamento	p. 12
3.1. Periodo di registrazione delle spese	p. 13
3.2. Stratificazione delle unità di primo stadio	p. 14
3.3. Schemi di rotazione delle unità di primo stadio	p. 18
3.4. La dimensione del campione	p. 24
4. Gli stimatori	p. 26
5. Alcune simulazioni basate sui dati dell'indagine del 1989	p. 29
5.1. Stimatori utilizzati nella simulazione	p. 29
5.2. L'importanza dell'ordine di inclusione delle unità di primo stadio nel campione	p. 32
5.2.1. Inclusione delle UPS in ordine decrescente	p. 32
5.2.2. Inclusione per prima della UPS più grande, e le altre in ordine crescente	p. 33
5.2.3. Inclusione delle UPS in ordine crescente	p. 33
5.2.4. Risultati corrispondenti ai diversi criteri di inclusione	p. 34
5.3. Valutazione del piano di campionamento proposto	p. 35
6. Conclusioni	p. 41
Riferimenti bibliografici	p. 43
Appendice A: tavole	p. 46
Appendice B: grafici	p. 62

1. Introduzione(*)

L'indagine sui consumi e sui bilanci delle famiglie, condotta dall'Istat, presenta difficoltà di vario tipo che inducono a presumere che l'errore extracampionario sia elevato. Infatti, il numero totale di mancate collaborazioni da parte dei Comuni (le unità di primo stadio o UPS) supera il 10% e rappresenta circa il 78% delle interviste non somministrate alle famiglie (unità di secondo stadio o USS), si veda Lucev (1992a). I rilevatori impiegati non hanno, in genere, una professionalità specifica e la supervisione della qualità del loro lavoro è carente. Inoltre, come accade in molte indagini campionarie, anche i rispondenti sono fonte di errore (con informazioni non appropriate, errori di memoria, sviste di registrazione, deliberata falsificazione dei dati) e tali errori sono amplificati sia dal tipo di modulo o diario utilizzato dalle famiglie nella registrazione delle spese, sia e soprattutto dalla scarsa preparazione dei rilevatori (Filippucci e Ferrante 1993; Filippucci e Zanelli 1994). Infine, la registrazione, la trascrizione dai moduli cartacei ai supporti magnetici, e l'imputazione durante la revisione automatica sono fonte di errori inevitabili (Filippucci e Marliani 1992a). Intervenire sulle mancate risposte per ridurre il loro ammontare non è né facile, né economico, anche se è possibile comunque utilizzare metodi adatti a migliorare le stime obiettivo (Falorsi e Russo 1992). Inoltre, alcuni errori possono essere rilevati e corretti con opportune metodologie sia in fase di predisposizione dei dati raccolti, sia in fase di elaborazione (Zieschang 1990; Lucev 1992b; Särndal 1992; Efron 1994). Lo sforzo maggiore, pertanto, deve essere indirizzato proprio a raccogliere le informazioni con più efficacia e precisione.

Per migliorare la tipologia e la qualità dei dati raccolti, la revisione dell'indagine sui consumi e sui bilanci delle famiglie deve interessare tutte le sue fasi e, in particolare, la strategia di rilevazione che coinvolge l'organizzazione operativa, la formazione del personale per qualificarne l'azione sul campo, gli strumenti di raccolta dei dati, e così via (Liepins e Uppuluri 1990; Filippucci e Marliani 1992b). Anche il piano di campionamento svolge un ruolo importante perché i risultati finali

(*) Si ringrazia Piero D. Falorsi per la lettura accurata del presente lavoro e per le utili osservazioni che hanno contribuito a migliorare l'esposizione di alcuni punti.

dell'indagine possono essere migliorati notevolmente approntando uno schema idoneo a conseguire le stime con la precisione desiderata e a ottimizzare le strategie di controllo e di acquisizione dei dati.

Nel presente lavoro vengono considerate alcune semplici variazioni al piano di campionamento attuale che permettano di ridurre le difficoltà sopra elencate e siano realizzabili in breve tempo. Tali modifiche sono attuabili facilmente e consentono una diminuzione sia delle UPS, sia delle USS costituenti il campione; inoltre, non comportano un notevole aumento dell'errore campionario. La riduzione delle unità, sia di primo sia di secondo stadio, è una condizione indispensabile per aumentare "l'efficacia" e il "rendimento" dell'indagine perché al crescere della dimensione del campione la spesa necessaria per ottenere una buona "qualità" dei dati diventa sempre più proibitiva e il processo di rilevazione (coordinamento, controllo, revisione, ecc.) incontra difficoltà organizzative notevoli per l'elevata quantità di lavoro richiesto. Infatti, a parità di massa globale di dati raccolti, semplicemente aumentando il periodo di registrazione delle spese per famiglia e mantenendo costante la dimensione del campione per periodo di rilevazione, il numero di USS in un trimestre o in un anno potrebbero diminuire. La diminuzione delle UPS e delle USS permette di controllare più agevolmente e con costi minori la qualità dei dati. Se si allunga, però, il periodo di registrazione delle spese c'è il rischio che il carico di lavoro degli intervistati influenzi negativamente l'attendibilità delle informazioni raccolte, ossia può intervenire un effetto "stanchezza" o "affaticamento" che induce l'intervistato a non rispondere accuratamente; ma ciò riguarda solo i beni che si riportano nel libretto per la rilevazione delle spese correnti e non i beni (durevoli) che vengono rilevati con l'intervista mensile (si veda il paragrafo 3.1).

I cambiamenti che vengono proposti potrebbero essere attuati per un periodo di transizione in cui l'Istat prepara un nuovo piano di campionamento basato su una stratificazione di UPS differenti dagli attuali Comuni, ossia aree geografiche definite sulla base dell'omogeneità rispetto alla spesa delle famiglie. Infatti, un campionamento stratificato di UPS areali sarebbe presumibilmente più efficace dell'attuale (Dippo e Jacobs 1983; Wolter e Causey 1991). Tuttavia, l'organizzazione amministrativa pone notevoli difficoltà all'attuazione di un campionamento areale di questo tipo. Inoltre, la creazione di una mappa

geografica sulla base di aree omogenee di spesa e l'aggiornamento della lista delle famiglie per ciascun area richiede un impegno notevole di risorse umane e finanziarie.

La struttura dell'esposizione è la seguente. Nel paragrafo 2 si descrive il disegno campionario attuale che costituisce il punto di riferimento per le modifiche proposte e si riportano alcune valutazioni sintetiche dell'indagine sui consumi dell'Istat del 1989 per delineare le caratteristiche essenziali del campione e delle grandezze oggetto di analisi. Nel paragrafo 3 si illustrano alcune possibili proposte di cambiamento inerenti al periodo di registrazione delle spese, alla stratificazione delle unità di primo stadio (i Comuni), agli schemi di rotazione delle UPS, e alla dimensione del campione rispetto alla precisione desiderata nelle stime dei parametri della popolazione. Nel paragrafo 4 si considerano gli stimatori sia in generale, sia con particolare riferimento alle possibili stime che si possono conseguire con i dati disponibili sui consumi rilevati dall'Istat per il 1989 perché tali stimatori saranno utilizzati per saggiare il piano di campionamento proposto tramite una simulazione. Nel paragrafo 5 si espongono i risultati di alcune simulazioni che, a partire dal campione attuale dell'Istat, generano una successione di campioni, con un numero diverso di UPS e con criteri differenti di selezione. In questo modo si è studiato il comportamento delle stime e dei relativi scarti quadratici medi all'aumentare delle UPS aggiunte al campione (generato) e si è potuto valutare l'efficienza relativa del piano di campionamento proposto rispetto a quello attualmente impiegato. Nel paragrafo 6 seguono le conclusioni.

2. Caratteristiche dell'indagine attuale sui consumi

La rilevazione dei consumi e dei bilanci di famiglia viene condotta dall'Istat mediante un campionamento a due stadi in cui le UPS sono i Comuni e le USS sono le famiglie. Il riferimento territoriale rispetto al quale viene garantita la precisione delle stime, sulla base delle spese distinte in diciannove gruppi, è la regione. In ciascuna regione si individuano due domini territoriali: il dominio auto-rappresentativo, AR, costituito dai Comuni capoluogo di provincia (95 unità nel 1989) o con popolazione maggiore o uguale a 50.000 abitanti (53 unità); il dominio

non auto-rappresentativo, NAR, costituito dai Comuni rimanenti (7936 unità nel 1989).

Ogni unità AR costituisce uno strato e il numero totale di strati AR è circa 150, che può variare leggermente da un anno all'altro. I Comuni AR partecipano all'indagine tutti i mesi dell'anno.

I Comuni NAR sono stratificati nell'ambito della regione geografica, in base all'altezza sul livello del mare (montagna, collina, pianura) e in base all'attività economica prevalente (agricoltura, industria, altre attività). Risultano così 135 strati NAR. All'interno di ogni strato NAR vengono selezionati tre Comuni e ognuno di essi rappresenta un gruppo: A, B, C. Ogni comune NAR partecipa all'indagine quattro mesi all'anno: il gruppo A partecipa i mesi di gennaio, aprile, luglio, e ottobre; il gruppo B partecipa nei mesi di febbraio, maggio, agosto, e novembre; il gruppo C partecipa nei mesi di marzo, giugno, settembre, e dicembre.

Il numero complessivo di Comuni campione in un mese è 285 ($150 + 135$); in un anno, come pure in un trimestre, è 555 ($150 + 135 \times 3$).

La selezione delle famiglie, in ciascun comune campione, viene attuata con il criterio sistematico utilizzando le liste anagrafiche dei Comuni stessi. Nei Comuni AR si realizzano dodici campioni distinti di famiglie, uno per ogni mese; mentre nei Comuni NAR si realizzano quattro campioni distinti di famiglie che effettueranno la rilevazione nei mesi relativi al gruppo a cui il comune appartiene. Ogni campione (mensile) delle famiglie viene suddiviso in tre sub-campioni: uno per ogni decade del mese, poiché la decade è il periodo di registrazione delle spese correnti. Il numero complessivo di famiglie campione nel mese è 3.300 circa (Innocenzi 1992; Falorsi e Russo 1992).

2.1. Alcune sintesi dell'indagine sui consumi del 1989

L'elaborazione dell'indagine Istat relativa a un anno (il 1989) è stata effettuata per evidenziare alcuni aspetti del piano di campionamento e delle relazioni esistenti tra le grandezze "tipiche" rilevate. In particolare, si sono esaminate tre voci di spesa per famiglia: la spesa totale, la spesa totale per beni alimentari e la spesa totale per beni non alimentari.

La tavola A1 dell'appendice A riporta il numero di famiglie intervistate per regione, per decade e per mese; inoltre, per ogni regione

riporta il numero medio di famiglie intervistate nell'anno (la media delle dimensioni dei campioni nelle 36 decadi), e il numero medio di famiglie per tipo di decade nell'anno (la media delle dimensioni dei campioni nelle 12 decadi simili). La tavola A2 riporta, a sua volta, il numero di famiglie intervistate per livello altimetrico e per attività economica prevalente. La dimensione effettiva del "campione decadale" è uguale a 939 famiglie, in media. Le dimensioni relative a ogni regione riflettono solo approssimativamente le differenze dell'ammontare delle USS regionali. La Sardegna, per esempio, presenta una dimensione campionaria che è circa uguale a quella dell'Abruzzo, del Friuli-Venezia-Giulia, dell'Umbria, della Basilicata, del Molise, e della Valle d'Aosta; ma la differenza tra esse è notevole rispetto al numero di Comuni nella regione (si veda la tavola 3 nel paragrafo 3.3 che riporta il numero di Comuni suddivisi in base all'appartenenza a montagna, collina e pianura), al coefficiente di variazione della spesa totale, al numero di abitanti, e al numero di famiglie (si veda la tavola 4 nel paragrafo 3.4). Infine, si può notare che le regioni Trentino-Alto-Adige, Emilia-Romagna, e Lombardia presentano una variabilità maggiore nelle ampiezze dei "campioni decadali" perché spesso la consegna delle interviste non avviene in tempi utili (Lucev 1992a).

La tavola A3 nell'appendice A riporta i coefficienti di variazione delle voci di spesa esaminate, per regione e per trimestre. La spesa per il consumo di beni alimentari mostra una certa costanza nel tempo e risulta pressoché uguale tra le diverse regioni, come era prevedibile; mentre la spesa per il consumo di beni non alimentari mostra un'alta variabilità intertemporale e interregionale. Pertanto, la variabilità del coefficiente di variazione della spesa totale è dovuta essenzialmente alla variabilità del coefficiente di variazione della spesa per il consumo di beni non alimentari. A questo riguardo, le regioni che, nell'anno, mostrano una maggiore variabilità sono Puglia, Basilicata, e Calabria.

Il coefficiente di variazione è stato analizzato in funzione della dimensione delle UPS per le 3 voci di spesa esaminate (si vedano i grafici da 1 a 6 nell'appendice B). L'enorme variabilità delle ampiezze delle UPS non mostra con sufficiente chiarezza la disposizione dei punti nel grafico; pertanto, ogni figura è stata riportata in due versioni: una completa, in cui sono incluse tutte le UPS; e l'altra incompleta, in cui sono incluse soltanto le UPS con un numero inferiore a 200.000 abitanti.

Dall'esame dei grafici si evince che la variabilità del coefficiente di variazione diminuisce al crescere dell'ampiezza delle UPS. In particolare, le UPS con un numero di abitanti superiore a 300.000 mostrano un coefficiente di variazione molto vicino a quello medio; mentre le UPS con un'ampiezza inferiore a 300.000 e superiore a 100.000 abitanti mostrano una certa variabilità, che aumenta notevolmente per quelle comprese tra 50.000 e 100.000 abitanti, e diventa ancora più grande per quelle con un'ampiezza inferiore a 50.000 abitanti. Naturalmente, queste soglie sono approssimative, ma possono costituire un elemento interessante per una eventuale stratificazione diversa da quella attuale, che verrà proposta nel paragrafo 3.2.

2.1.1. Alcune analisi della varianza

Le tavole A4, A5, e A6 nell'appendice A riportano i livelli medi di spesa e i relativi scarti quadratici medi per regione e per trimestre, suddivisi tra AR e NAR. I dati delle tavole si riferiscono al campione rilevato e sono stati pesati con i coefficienti di riporto trimestrali, in modo che le medie e le deviazioni standard risultino dello stesso ordine di grandezza delle medie e delle deviazioni standard non pesate. Le tavole forniscono l'ordine di grandezza dei caratteri esaminati rispetto ai fattori considerati; inoltre, costituiscono una base descrittiva per una serie di applicazioni dell'analisi della varianza rispetto a diversi fattori che hanno permesso di comprendere meglio la natura della variabilità del fenomeno in esame.

Fattore AR-NAR. L'analisi della varianza delle tre voci di spesa a un criterio di classificazione mostra che non c'è differenza significativa tra le unità AR e NAR nei livelli medi di spesa per il consumo totale e per il consumo di beni non alimentari; mentre vi è una differenza significativa tra la spesa per il consumo di beni alimentari ($F_{1,33388} = 10,67; p < .005$). I valori medi, in migliaia, sono 555 per le unità AR ($n=16403$) e 543 per le unità NAR ($n=16987$), ma la deviazione standard risulta contenuta: 432.6 per le unità AR e 541.1 per le unità NAR sicché la differenza diventa significativa.

Fattori: AR-NAR e trimestre. L'analisi della varianza relativa alle tre voci di spesa a due criteri di classificazione mostra che la spesa per il

consumo totale e per il consumo di beni non alimentari non presentano interazioni significative, né differenza tra le unità AR e NAR. La differenza tra i trimestri, invece, è molto marcata sia nella spesa per il consumo totale ($F_{1,33660} = 57,24; p < .001$), sia nella spesa per il consumo di beni non alimentari ($F_{1,33660} = 44,30; p < .001$). La spesa per il consumo di beni alimentari mostra, invece, una interazione significativa e anche differenze significative tra i livelli medi di spesa sia tra le unità AR e NAR, sia tra i trimestri; ma, dato che l'interazione è significativa, questi ultimi risultati dovrebbero essere analizzati con maggiore attenzione qualora si fosse interessati a comprendere dove si verificano le differenze (Lindman 1991).

Fattori: regione e AR-NAR. L'analisi della varianza relativa alle tre voci di spesa a due criteri di classificazione mostra che tutte presentano interazioni significative tra i fattori *regione* e *AR-NAR*; risultano significative anche le differenze tra i livelli medi di spesa delle regioni e del fattore AR-NAR. In particolare, le regioni in cui si osservano le differenze significative tra le unità AR e NAR, nei livelli medi della spesa totale per il consumo e per il consumo di beni non alimentari, sono: Piemonte ($p < .001$), Lombardia ($p < .001$), Trentino-Alto-Adige ($p < .044$), Veneto ($p < .001$), Liguria ($p < .026$), Toscana ($p < .036$), Marche ($p < .002$), Lazio ($p < .001$), Molise ($p < .007$), Sicilia ($p < .001$), e Sardegna ($p < .001$). La differenza tra le regioni è ancora più marcata nella spesa per il consumo di beni alimentari: rispetto all'elenco precedente si elimina il Trentino-Alto-Adige e si aggiungono Emilia-Romagna, Umbria, Campania, Puglia, e Calabria.

Fattori: regione, AR-NAR, trimestre. L'analisi della varianza relativa alla tre voci di spesa a tre criteri di classificazione mostra che l'interazione tra AR-NAR e trimestre non è significativa, mentre l'interazione tra i tre i fattori si rivela significativa solo nella spesa per il consumo di generi alimentari. Risultano significative le interazioni regione-(AR-NAR) e regione-trimestre, come pure le differenze tra i livelli medi di spesa, anche se per il fattore AR-NAR il livello di significatività è solo del 5%.

L'analisi della varianza rispetto ai fattori (AR-NAR), regione, e trimestre ha mostrato, quindi, che il fattore AR-NAR svolge un ruolo complesso perché a volte presenta livelli di spesa statisticamente differenti e altre volte, invece, le differenze a esso imputabili non sono



rilevanti. Tuttavia, si può concludere che non si tratta di un elemento trascurabile nell'individuazione di strati omogenei di spesa e ciò equivale a sostenere che la stratificazione per dimensione delle UPS è importante.

2.1.2. Un'analisi dei gruppi

Si è eseguita, infine, un'analisi dei gruppi (*cluster-analysis*) il cui fine non era il raggruppamento delle regioni italiane per ridurre il numero di UPS nel campione poiché, come si è detto, la precisione delle stime è ancorata alla suddivisione del territorio in regioni, ma l'accertamento della loro "similitudine" rispetto alla tre voci di spesa per esplorare e verificare la potenziale possibilità di aggregarle. Nel grafico 7 dell'appendice B è riportato il grafico (dendrogramma) che illustra il procedimento di aggregazione delle unità (regioni) ottenuto con il metodo di Ward, fissando il numero dei gruppi finali a cinque (Ricolfi 1992) e utilizzando la spesa media per il consumo totale, la relativa deviazione standard, e la differenza tra le spese per il consumo non alimentare e le spese per il consumo alimentare. L'utilizzo di tali caratteristiche permette di tener conto del livello medio della spesa totale per regione, della sua variabilità, e del differente livello medio esistente tra le spese per il consumo alimentare e non alimentare. Il risultato ottenuto nel grafico 7 ha mostrato una sorprendente stabilità sia rispetto ad altri metodi di raggruppamento e sia rispetto ai tipi di spesa utilizzati; in particolare, i seguenti gruppi sono risultati più "omogenei" tra loro: (Lombardia, Trentino-Alto-Adige, Veneto, e Emilia-Romagna), (Piemonte, Liguria, Abruzzo), (Friuli-Venezia-Giulia, Marche, Toscana, e Sardegna). Le altre regioni, invece, hanno mostrato maggiore facilità sia a formare gruppi tra loro, sia ad aggregarsi ai gruppi precedenti, sia a separarsi e a costituire gruppi più piccoli.

3. Proposte di modifiche al piano di campionamento

I piani di campionamento per la rilevazione dei consumi e dei bilanci delle famiglie, come anche la struttura e la strategia di indagine, variano da paese a paese: per una recente rassegna, si veda Innocenzi (1992). Nei

paragrafi successivi si esaminano alcuni elementi che caratterizzano e influenzano il piano di campionamento: il periodo di rilevazione delle spese, che varia da una settimana a un mese, e a volte non è costante per diverse categorie di beni (paragrafo 3.1), la distinzione tra città grandi e piccole o una stratificazione territoriale (paragrafo 3.2), lo schema di rotazione degli elementi costituenti gli strati o UPS (paragrafo 3.3), e l'ampiezza del campione (paragrafo 3.4).

3.1. Periodo di registrazione delle spese

La registrazione giornaliera delle spese correnti viene effettuata per un periodo di tempo la cui ampiezza non è determinata in base a regole precise e varia da paese a paese (10, 14, 30 giorni o 1 mese). L'EUROSTAT ha emanato, perciò, delle direttive rivolte ai paesi della Comunità europea, che fissano un periodo di registrazione delle spese correnti pari a 14 giorni e coincidenti con due settimane. Se nell'indagine attuale dell'Istat si introducesse la *quattordicina* si otterrebbe un effetto positivo che consisterebbe nella riduzione del numero di periodi di registrazione all'anno da 36 (decadi) a 26 (quattordicine) e nella riduzione del 27-28% del numero totale di famiglie incluse nel campione, cioè da circa 39000 a circa 25000 famiglie all'anno. Tale dimensione è confrontabile con quella riportata nella tavola 4, ottenuta con un errore relativo campionario desiderato pari al 22% per periodo di rilevazione. La precisione che ne deriva è sicuramente più alta nel caso di una variabilità poco accentuata tra le quattordicine, che facilita la cumulabilità dei campioni bisettimanali per "riportare" la situazione familiare a base trimestrale o annua. Naturalmente, ciò riguarda solo i beni (non durevoli) che vengono registrati nel "Libretto per la rilevazione delle spese correnti" (Mod. Istat /W/405) perché i beni durevoli, invece, vengono rilevati con l'intervista mensile e registrati nel questionario "Il riepilogo delle spese" (Mod. Istat /W/406) (Innocenzi 1992).

Un'obiezione rilevante alla quattordicina concerne l'eccessiva lunghezza del periodo di registrazione perché potrebbe provocare un aumento del numero di famiglie che non partecipano all'indagine e indurre una maggiore inattendibilità dei dati raccolti a causa della stanchezza degli intervistati che li scoraggia a fornire informazioni

accurate. Infatti, l'indagine sui consumi richiede un notevole impegno per la considerevole quantità di elementi da riportare nei questionari. Gli effetti del passaggio dalla decade alla quattordicina sugli intervistati potrebbero essere valutati con un piano sperimentale per verificare se le difficoltà menzionate siano più o meno rilevanti rispetto ai vantaggi che ne deriverebbero.

La scelta di due settimane di rilevazione per la stessa famiglia pare quella ottimale perché le abitudini di spesa dei consumatori sembrano "sincronizzate" con la settimana e i modelli di spesa mostrano incoerenza con la decade (Filippucci e Marliani 1992a): lo sfasamento dell'indagine attuale rispetto alla settimana non permette di rilevare (sulla stessa unità) questa ciclicità. La stima dei consumi per trimestre richiederebbe, però, calcoli di imputazione perché non c'è coincidenza tra settimane e trimestre (si noti che ciò accade, però, anche con la quattordicina). Per ottenere tale coincidenza si potrebbe assumere o un periodo di registrazione mensile, come in Germania, Svizzera, Danimarca, Lussemburgo, Paesi Bassi; o quindicinale, per esempio dall'inizio alla metà del mese e dal giorno successivo alla metà fino alla fine del mese. Tuttavia, nel caso mensile, la mole di lavoro elevata richiesta alle famiglie provoca una "caduta" notevole delle risposte; nel caso quindicinale, il secondo periodo risulterebbe lungo 16 giorni nei mesi che hanno 31 giorni e 14 o 13 in febbraio; in ambedue i casi, si ha variabilità del periodo di riferimento delle registrazioni delle spese e "sfasamento" con le settimane che possono incidere, invece, sui modelli di spesa.

3.2. *Stratificazione delle unità di primo stadio*

Il raggruppamento delle UPS in aree territoriali omogenee rispetto alla tipologia e alle modalità della spesa renderebbe, ovviamente, più efficace la stratificazione rispetto a quella attuale. Sarebbe pertanto opportuno condurre un'indagine per individuare le unità territoriali in cui si abbia uniformità di comportamenti dei consumatori e stratificare in base a tali unità per migliorare le stime (Filippucci e Marliani 1992a,b). La scelta del metodo del campionamento areale, tuttavia, richiederebbe un periodo di tempo abbastanza lungo prima di poter essere attuato e l'Istat non solo dovrebbe affrontare uno sforzo notevole per costruire e

aggiornare le aree omogenee di spesa, ma dovrebbe anche gestire, per ogni area, i corrispondenti registri anagrafici per selezionare le USS. Si potrebbero campionare le abitazioni invece delle famiglie, ma in tal caso bisognerebbe costruire una lista delle aree abitative e diventerebbe più complesso conseguire le stime. Esaminiamo, quindi, l'impatto di una stratificazione delle unità di primo stadio, che si baserà ancora sulla distinzione tra Comuni AR e NAR (come nel piano di campionamento attuale) fino a quando non sarà disponibile una suddivisione alternativa del territorio. D'altronde, anche l'EUROSTAT, precisando le variabili e gli schemi che intende pubblicare per l'indagine campionaria sui consumi delle famiglie, ha definito la tipologia dei Comuni in base alla loro densità demografica (Innocenzi 1992).

In questo paragrafo si propone una stratificazione delle UPS diversa da quella attuale suggerita da alcuni risultati emersi dall'elaborazione dei dati dell'indagine sui consumi del 1989 (si vedano i paragrafi 2.1 e 5.2; Filippucci e Marliani 1992a,b; Drudi 1992). I risultati più significativi per le proposte che seguono sono stati riassunti nei tre punti elencati di seguito.

1. La stima della spesa totale e la relativa varianza campionaria e non cambiano significativamente con la diminuzione del numero di unità di primo stadio, ossia mostrano una sorprendente stabilità. Per corroborare la proposta di riduzione delle UPS, è stata eseguita una simulazione in cui parte delle UPS venivano eliminate casualmente per ogni regione. Più specificamente, in ogni regione si eliminavano almeno una UPS di tipo AR non capoluogo di regione e blocchi di UPS di tipo NAR non superiori al 30% del totale. Il confronto dei risultati del campione ottenuto con la simulazione e del campione completo originario non ha mostrato differenze significative dei livelli di spesa in ambito regionale. Pertanto, una riduzione delle UPS (inferiore al 30% delle attuali) non dovrebbe influenzare marcatamente i risultati delle stime (si veda anche il paragrafo 5).

2. Si deve notare che non esistono prove empiriche sufficienti per giustificare la scelta della soglia di 50.000 abitanti per distinguere i Comuni NAR dagli AR. Per esempio, il coefficiente di variazione per diversi tipi di spese, espresso in funzione dell'ampiezza del Comune di appartenenza delle famiglie, è altamente fluttuante nelle ampiezze piccole e diviene stabile solo dopo ampiezze superiori a 300.000 abitanti (si

riprendano i grafici da 1 a 6 riportati nell'appendice B). Tuttavia, l'alta variabilità tra le UPS di piccole dimensioni potrebbe derivare anche dal basso numero di famiglie selezionate nelle UPS con un'ampiezza demografica piccola. Infatti, il numero di famiglie in ogni UPS è spesso minore di 10 e, su tutto l'anno, le famiglie provenienti da unità primarie NAR sono meno di un terzo di quelle estratte dalle unità primarie AR. Pertanto, diminuendo il numero di UPS e aumentando il numero di famiglie (le USS) selezionate per ogni UPS si dovrebbe ottenere un miglioramento della strategia di rilevazione in termini di organizzazione e qualità dei dati raccolti. Si noti che un aumento del numero di famiglie per le unità NAR si può ottenere anche aumentando il periodo di permanenza dell'unità nel campione.

3. Non si sono riscontrate differenze significative né tra i livelli di spesa e né riguardo alla variabilità della spesa nei Comuni AR e NAR. Naturalmente, non ne consegue immediatamente che la stratificazione sull'ampiezza delle UPS non è determinante ai fini della stima perché l'analisi della varianza ha mostrato che tale fattore svolge un ruolo complesso, che varia per regione e per trimestre di rilevazione, e in molti casi la differenza tra i livelli medi di spesa osservati nelle unità AR e NAR risultano significativamente differenti.

In base a queste considerazioni, una possibile stratificazione delle UPS, basata su quattro tipi di strati, è la seguente.

La classe 4 contiene le unità AR ridefinite, ossia i Comuni con un numero di abitanti superiore o uguale a 300.000; così, le unità AR si riducono da 150 a 11 (l'elenco appare nella successiva tavola 4, nel paragrafo 3.4). La precisione delle stime per regione implica l'opportunità di avere almeno un Comune AR per ogni regione; così, per quelle che non hanno un Comune con un numero di abitanti superiore a 300.000 (la Valle d'Aosta, il Trentino-Alto-Adige, il Friuli-Venezia-Giulia, le Marche, l'Umbria, l'Abruzzo, il Molise, la Basilicata, la Calabria, e la Sardegna) si considera AR il Comune con il maggior numero di abitanti: Aosta, Trento, Trieste, Ancona, Perugia, Pescara, Campobasso, Potenza, Reggio di Calabria, e Cagliari; a eccezione dell'Abruzzo, coincidono con il capoluogo di regione. Quindi, il numero di Comuni AR diventa 21.

La classe 3 contiene i Comuni che hanno un'ampiezza da 100.000 a 299.999 abitanti (l'elenco appare nella successiva tavola 1, nel paragrafo

3.3). I membri di questa classe sono attualmente 32. Al fine di avere un multiplo di tre (33), utile per gli schemi di rotazione, si è aggiunto il Comune di Udine che al 31 dicembre 1989 contava 98.872 abitanti.

La classe 2 contiene i Comuni che hanno un'ampiezza da 50.000 a 99.999 abitanti (l'elenco appare nella tavola 2, nel paragrafo 3.3). I membri di questa classe sono attualmente 83, ma si noti che questa è la classe che può risentire maggiormente delle variazioni demografiche perché vi sono diversi Comuni con un numero di abitanti molto vicino alla soglia inferiore dello strato. Mazara del Vallo è stata aggiunta per ottenere un multiplo di 3 e facilitare gli eventuali schemi di rotazione.

La classe 1 contiene i rimanenti Comuni che hanno un'ampiezza inferiore a 50.000 abitanti. Per questa classe, che include gli attuali Comuni NAR, si può mantenere la stratificazione in base al livello altimetrico dei Comuni stessi (montagna, collina, e pianura) perché tale carattere sembra mostrare una certa influenza sul tipo e sul livello di spesa (Falorsi *et al.* 1992). Si noti, comunque, che la stratificazione altimetrica implicitamente tiene conto dell'ampiezza demografica dei Comuni che tende a decrescere al crescere dell'altezza sul livello del mare. Si sta verificando, però, una continua omologazione del comportamento dei consumatori, dovuto ai mezzi di comunicazione di massa, tanto che si può supporre che la stratificazione in base all'attività economica prevalente ottenuta con il censimento (agricoltura, industria, altre attività) non influenzi significativamente il tipo e il livello di spesa.

La stratificazione proposta annulla la soglia di 50.000 abitanti portandola a 300.000 abitanti; inoltre, può essere applicata facilmente e produce una riduzione delle UPS attraverso schemi di rotazione delle unità appartenenti alle classi 3 e 2 che contengono proprio le attuali unità AR che nello schema proposto non dovrebbero più essere AR. Tale soluzione è operativamente facile, rispetta condizioni di ragionevolezza, ma non si propone come una stratificazione "ottima" perché non rispetta le condizioni usuali, per esempio, la costruzione di strati aventi ampiezza approssimativamente costanti in termini di unità ivi residenti. Gli estremi delle ampiezze delle UPS che determinano gli strati sono stati derivati esclusivamente sulla base di alcuni risultati osservati nel campione dell'indagine del 1989 e, quindi, "dipendono" da questa. Infatti, tale procedimento vuole rappresentare, come si è detto, solo un punto di partenza per stimolare la discussione e l'approfondimento del problema.

3.3. Schemi di rotazione delle unità di primo stadio

L'introduzione di una stratificazione per ampiezza demografica dei Comuni che sono attualmente AR implica che, a ogni periodo di rilevazione, nell'indagine modificata secondo lo schema proposto sarà presente un numero inferiore di UPS ex-AR, rispetto all'indagine corrente. La scelta dei Comuni da inserire nel campione per i diversi strati può essere determinata casualmente sulla base di un numero fissato di UPS da selezionare in ogni strato al fine di facilitare l'organizzazione dell'indagine (caso A). Un'alternativa è costituita da uno schema di rotazione fisso che consente a tutti i Comuni di partecipare secondo un "calendario" determinato perché si facilita la conduzione dell'indagine e si è certi che, seppure a turno, contribuiscono tutti (o quasi) al conseguimento delle stime (caso B). In entrambi i casi sorgono difficoltà inerenti ai domini di stima regionali, ma diversamente si avrebbe che le UPS delle classi 2 o 3 diventerebbero di fatto AR nelle regioni che ne hanno una sola in ognuna di esse. La stabilità delle stime, che si ottiene con le UPS AR del piano di campionamento modificato (si veda il capitolo 5), sembra indicare che tali problemi, anche se rilevanti, non dovrebbero incidere notevolmente sui risultati conseguiti con l'indagine.

L'introduzione della rotazione delle unità di primo stadio, appartenenti alle classi 2 e 3, è un espediente che permette in realtà di censire le unità dello strato in un arco di tempo relativamente breve. Tuttavia, la rotazione implica ugualmente un aggravio per l'Istat, che periodicamente dovrebbe allacciare contatti con Comuni sempre diversi. Di fatto, i contatti con gli attuali Comuni AR verrebbero scaglionati secondo gli schemi di rotazione che fissano le modalità di inclusione delle UPS nell'indagine e il periodo di permanenza nel campione.

I Comuni della classe 3 potrebbero essere suddivisi in tre gruppi (a , b , c) e inseriti nel campione con una quota fissa pari a due terzi nel primo semestre ($a+b$); nel semestre successivo, il gruppo (a) viene escluso e il rimanente gruppo (c) viene aggiunto al campione; nel semestre successivo ancora, il gruppo (b), cioè quello che è stato più tempo nel campione, viene escluso e rientra quello escluso in precedenza (a). A regime, perciò, un terzo delle unità statistiche di primo stadio rimane nel campione per un periodo di 12 mesi, cioè un anno, mentre gli altri due terzi rimangono 6 mesi ognuno nel "campione annuo". Se si indicano con

a , b , c i tre gruppi e si riferisce la rotazione all'anno solare, t , lo schema proposto si può rappresentare come segue:

$$[(a+b)_{I \text{ sem.}}, (b+c)_{II \text{ sem.}}]_t \rightarrow [(c+a)_{I \text{ sem.}}, (a+b)_{II \text{ sem.}}]_{t+1} \rightarrow \\ [(b+c)_{I \text{ sem.}}, (c+a)_{II \text{ sem.}}]_{t+2} \rightarrow [(a+b)_{I \text{ sem.}}, (b+c)_{II \text{ sem.}}]_{t+3} \rightarrow \dots$$

Nella tavola 1 è riportata una possibile suddivisione in tre gruppi dei Comuni della classe 3. La scelta dei Comuni per formare i gruppi di rotazione potrebbe essere effettuata su tutto il paese casualmente perché le unità appartenenti allo strato sono omogenee tra loro, rispetto al carattere di stratificazione. Tuttavia, la costituzione dei gruppi è stata effettuata casualmente all'interno di ogni regione perché il riferimento territoriale regionale della precisione delle stime è sempre molto importante. Le regioni che non avevano un numero sufficiente di UPS per formare un "gruppo completo di rotazione" (cioè, un multiplo di tre) sono state aggregate, quindi, alle regioni più vicine o simili. In base allo schema sopra riportato, per un dato anno, i Comuni elencati nelle colonne indicate con le lettere a e b della tavola 1 restano nel campione il primo semestre; nel secondo semestre si sostituiscono i Comuni della colonna a con quelli della colonna c , così lo schema sarà ($a+b$) nel primo semestre e ($b+c$) nel secondo; pertanto i Comuni della colonna b restano un anno intero nel campione. Nell'anno successivo "rientrano" i Comuni della colonna a al posto di quelli della colonna b restando per tutto l'anno, quelli della colonna c restano i primi sei mesi, e quelli della colonna b "rientrano" e restano nel secondo semestre; lo schema sarà ($c+a$) nel primo semestre e ($a+b$) nel secondo. Negli anni seguenti si procede con lo stesso meccanismo, come specificato nello schema precedente. In tal modo quattro regioni (Lombardia, Veneto, Toscana, e Puglia) presentano una rotazione equilibrata (perfetta); mentre l'Emilia-Romagna presenta sempre UPS in ogni gruppo, ma non sono sempre lo stesso numero (2 nel gruppo a , 3 nel gruppo b , 3 nel gruppo c). La conseguenza, però, è che per ottenere le stime annuali bisogna pesare opportunamente le stime trimestrali.

Tavola 1 — Comuni appartenenti alla classe 3 (100.000–299.999 abitanti) e relativa popolazione residente al 31 Dicembre 1989, divisi in 3 gruppi.

<i>a</i>		<i>b</i>		<i>c</i>	
Città	Abitanti	Città	Abitanti	Città	Abitanti
Novara	103.088	Bolzano	100.707	Udine(*)	98.872
Monza	123.073	Bergamo	117.584	Brescia	196.535
Verona	258.476	Vicenza	109.109	Padova	220.358
La Spezia	104.511	Parma	174.341	Piacenza	104.023
Ferrara	141.404	Modena	176.857	Reggio Emilia	130.825
Ravenna	136.166	Forlì	109.986	Rimini	130.638
Pisa	102.150	Livorno	171.346	Prato	165.888
Torre d. Greco	103.577	Latina	101.936	Terni	110.020
Taranto	244.512	Foggia	159.199	Lecce	101.957
Cosenza	105.349	Catanzaro	103.521	Salerno	152.159
Messina	273.570	Siracusa	124.606	Sassari	119.717

(*) Udine è stata inclusa per ottenere un multiplo di 3.

I Comuni della classe 2 potrebbero essere suddivisi anch'essi in tre gruppi (*a*, *b*, *c*) e inseriti nel campione con una quota fissa pari a un terzo nel primo semestre dell'anno (*a*), un terzo nel secondo semestre dell'anno (*b*), un terzo nel primo semestre dell'anno successivo (*c*), sicché un Comune "ruota" nell'anno e non figura sempre nello stesso semestre. A regime, quindi, ogni UPS resta sei mesi nel campione:

$$(a, b)_t \rightarrow (c, a)_{t+1} \rightarrow (b, c)_{t+2} \rightarrow (a, b)_{t+3} \dots$$

Nella tavola 2 sono elencati i Comuni della classe 2 in base a un possibile schema di rotazione che tiene conto, come nel caso precedente, della vicinanza o similarità delle regioni tra loro. Nel primo semestre sono inclusi nel campione solo i Comuni elencati nella colonna indicata con la lettera *a*, nel secondo solo i Comuni elencati nella colonna indicata con la lettera *b*: fase (*a*, *b*). Nell'anno successivo si includono nel campione solo i Comuni della colonna *c* nel primo semestre e solo quelli della colonna *a* nel secondo: fase (*c*, *a*). Si procede così, di volta in volta, come indicato nello schema precedente.

Tavola 2 — Comuni appartenenti alla classe 2 (50.000–99.999 abitanti) e relativa popolazione residente al 31 Dicembre 1989, divisi in 3 gruppi (le linee evidenziano l'appartenenza delle UPS alle diverse regioni)

<i>a</i>		<i>b</i>		<i>c</i>	
Città	Abit.	Città	Abit.	Città	Abit.
Moncalieri	61.338	Rivoli	53.518	Biella	51.365
Vercelli	50.313	Cuneo	55.746	Asti	74.649
Alessandria	93.866	San Remo	59.635	Savona	69.806
Busto Arsizio	77.857	Cinisello Bals.	78.046	Cologno Monz.	52.838
Pavia	80.653	Rho	51.637	Sesto S. Giov.	89.517
Mantova	54.808	Vigevano	61.731	Cremona	75.547
Como	89.602	Varese	88.018	Pordenone	50.300
Chioggia	53.581	Treviso	84.066	Rovigo	52.498
Carpi	60.690	Imola	62.042	Faenza	54.118
Cesena	89.606	Carrara	68.528	Massa	67.570
Lucca	86.676	Viareggio	60.257	Pistoia	89.972
Scandicci	54.399	Arezzo	91.527	Siena	58.278
Grosseto	71.034	Fano	52.890	Foligno	53.563
Pesaro	90.325	Ascoli Piceno	52.667	Viterbo	59.798
Civitavecchia	51.094	Guidonia M.	56.052	Tivoli	54.352
L'Aquila	67.348	Teramo	52.501	Chieti	57.362
Benevento	64.842	Caserta	67.769	Aversa	57.817
Castellam. di St.	68.478	Casoria	79.729	Afragola	62.359
Ercolano	63.571	Giugliano in C.	57.041	Portici	73.488
Pozzuoli	76.121	Cava de' Tirreni	52.341	Torre Annunz.	56.471
S. Giorgio a Cr.	63.052	Avellino	55.886	Cerignola	54.549
Manfredonia	58.920	San Severo	55.017	Altamura	56.631
Andria	89.762	Barletta	88.074	Bitonto	53.292
Molfetta	64.215	Brindisi	92.815	Matera	54.377
Crotone	61.688	Lamezia Terme	68.985	Marsala	80.869
Trapani	72.837	Agrigento	56.372	Caltanissetta	62.588
Gela	79.601	Modica	50.556	Ragusa	68.850
Vittoria	55.870	Mazara d. V.(*)	49.330	Quartu S. Elena	62.042

(*) Mazara del Vallo è stata inclusa per ottenere un multiplo di 3.

I Comuni della classe 1 corrispondono agli attuali NAR e possono essere stratificati in base all'altezza sul livello del mare e suddivisi secondo l'appartenenza a montagna, pianura, e collina. Gli strati effettivi (tavola 3) risultano 51 perché Valle d'Aosta e Trentino-Alto-Adige hanno solo Comuni in montagna, mentre Liguria, Umbria, Marche, Abruzzo, e Molise non hanno Comuni in pianura. Nella colonna relativa alla classe 1

della tavola 3 è riportata la proposta del numero totale di Comuni da selezionare per regione, determinato in base al seguente criterio: si seleziona *una* UPS se il loro numero totale per strato è minore di 70, *due* UPS se il loro numero totale per strato è maggiore di o uguale a 70 e minore di 700, *tre* UPS se il loro numero totale per strato è maggiore di o uguale a 700. Le UPS così selezionate sono 90 perché per il Trentino Alto Adige si sono considerate separatamente Bolzano e Trento. Tuttavia, se si decidesse che l'organizzazione e la conduzione dell'indagine risulta più semplice con la selezione di un numero fisso di UPS per strato, allora si potrebbero selezionare *due* UPS per strato. Se si aggregasse lo strato "montagna" della Puglia, che contiene solo 8 UPS, allo strato "collina", si otterrebbero 50 strati effettivi e 100 UPS da selezionare.

I Comuni della classe 1 potrebbero rimanere nel campione per un semestre e nel semestre successivo essere sostituite con nuove unità. La proposta del semestre, per il periodo di permanenza delle unità della classe 1 nel campione, deriva da considerazioni logistiche per facilitare le rilevazioni da parte dei Comuni selezionati. Organizzare una rilevazione comporta l'impiego di personale che deve essere istruito e che, quando è fornito dalle amministrazioni locali, viene sottratto a altri compiti. Se il periodo di coinvolgimento dei Comuni è breve, è più probabile che subentri il disinteresse e la superficialità nella raccolta dei dati; inoltre, l'Istat può effettuare controlli con efficacia e sistematicità. Per esempio, in un semestre, nella Basilicata si avrebbero 4 Comuni della classe 1 che si possono controllare in parte o globalmente, nelle 13 quattordicine di permanenza. In questo caso, ha senso addestrare gli intervistatori e essi stessi avrebbero maggiore interesse a apprendere perché per un semestre avrebbero l'onere di contattare le famiglie per rilevare le informazioni pertinenti all'indagine sui consumi e sui bilanci. In condizioni di questo tipo, si possono mandare dei supervisori a effettuare le interviste quando gli addetti del Comune non eseguono i loro compiti oppure i dati sono incompleti o presentano numerosi errori e incongruenze. Potrebbe accadere che, per un Comune piccolo, il numero di famiglie residenti sia insufficiente a "coprire" l'intero periodo della permanenza nel campione. In tal caso, si potrebbe effettuare una selezione tra le famiglie già intervistate e, in caso di rifiuto, si potrebbe estrarre un altro Comune "simile", ossia con un numero di famiglie circa uguale a quello già intervistato, ottenendo perciò una sostituzione; alternatively, nella

classe 1 bisognerebbe applicare una stratificazione ulteriore secondo l'ampiezza delle UPS per discernere meglio (rispetto a questo fattore) il diverso comportamento delle famiglie nell'acquisto di beni.

Tavola 3 — Numero totale di UPS, numero di Comuni nella classe 1 per altezza sul livello del mare: montagna (M), collina (C), pianura (P). Massimo* numero di UPS nel campione per classe di ampiezza demografica delle UPS.

Regione	Strati per altitudine				n_i			Classi di amp.			
	T.	M.	C.	P.	M.	C.	P.	1*	2*	3*	4
Piemonte	1209	350	584	275	2	2	2	6	3	1	TO
Valle d'Aosta	74	74	—	—	2	—	—	2	—	—	AO
Lombardia	1546	479	319	748	2	2	3	7	4	2	MI
Trentino-Alto-Adige	339	339	—	—	4	—	—	4	—	1	TN
Veneto	582	119	124	339	2	2	2	6	2	2	VE
Friuli-Venezia-Giulia	219	59	51	109	1	1	2	4	—	1	TS
Liguria	235	113	122	—	2	2	—	4	1	1	GE
Emilia-Romagna	341	67	110	164	1	2	2	5	2	6	BO
Toscana	287	79	186	22	2	2	1	5	3	2	FI
Umbria	92	17	75	—	1	2	—	3	1	1	PG
Marche	246	48	198	—	1	2	—	3	1	—	AN
Lazio	375	119	243	13	2	2	1	5	2	1	RM
Abruzzo	305	165	140	—	2	2	—	4	1	—	PE
Molise	136	84	52	—	2	1	—	3	—	—	CB
Campania	549	128	317	104	2	2	2	6	5	2	NA
Puglia	257	8	65	184	1	1	2	4	3	2	BA
Basilicata	131	78	47	6	2	1	1	4	1	—	PZ
Calabria	409	153	234	22	2	2	1	5	1	2	RC
Sicilia	388	96	259	33	2	2	1	5	3	2	2+
Sardegna	366	38	254	74	1	2	2	5	1	1	CA
Totale	8086	2613	3380	2093	36	32	22	90	28	22	21

* Numero massimo perché in alcune fasi della rotazione possono diminuire di una unità.

1) 0-49.999 abitanti; 2) 50.000-99.999 a.; 3) 100.000-299.999 a.; 4) ≥ 300.000 a.

+ Palermo e Catania

Dalla tavola 3, le unità di primo stadio risulterebbero 161 al mese, contro le 285 del disegno attuale. In un anno solare, le 150 unità AR attuali diventano 110: 21 Comuni AR della classe 4, 33 Comuni della classe 3 dei quali 22 stanno un solo semestre (11 nel primo e 11 nel secondo), 56 Comuni della classe 2 (28 il primo semestre e 28 il secondo). Gli strati dei Comuni NAR della classe 1, corrispondente ai NAR del disegno attuale che sono 135, diventano 51; il numero di UPS

da estrarre è 180 l'anno (90 nel primo semestre e 90 nel secondo); il numero semestrale, 90, potrebbe essere ridotto a un minimo di 51: un Comune per ogni strato. Pertanto, il numero complessivo di Comuni campione, all'anno, sarebbe di 290 contro i 455 del disegno attuale. Si noti che una ulteriore riduzione delle unità di primo stadio si ottiene modificando il numero di Comuni proposto per la classe 1 nella tavola 3, che ovviamente è solo indicativo, e/o cambiando lo schema di rotazione per le unità delle classi 2 e 3.

3.4. La dimensione del campione

Si vuole dare, ora, qualche valutazione sulla ripartizione delle famiglie del campione (USS) nelle diverse regioni. Per ottenere un risultato utile in termini concreti bisognerebbe calcolare la dimensione del campione per strato al fine di rispondere alla domanda: "Quante famiglie bisogna selezionare dalla lista dell'anagrafe del tal Comune selezionato nel tal strato?". Tuttavia, in questo ambito non ha senso una proposta così dettagliata. Può essere comunque utile, invece, valutare i livelli di precisione delle stime comparando l'allocatione attuale per ciascuna regione con una possibile alternativa, che possa tenere conto delle informazioni già cumulate con l'esperienza svolta.

Nell'indagine sui consumi delle famiglie, l'Istat pubblica stime per ciascuna voce di spesa distinta in base a 19 gruppi e 60 categorie di consumo. Il riferimento territoriale della precisione delle stime sulla base dei gruppi, come si è detto, è la regione, mentre sulla base delle categorie è l'intero paese. La dimensione del campione, quindi, può essere determinata per regione e una valutazione approssimativa che tenga conto della variabilità del carattere nella regione (Cochran 1977, pp. 77-78) si può così ottenere con la seguente

$$n_r = \min \left\{ \max_{1 \leq k \leq 19} \frac{(C_{kr} / \gamma)^2}{1 + [(C_{kr} / \gamma)^2] / N_r}, n_r^* \right\}$$

dove C_{kr} è il coefficiente di variazione della k -esima classe di spesa nella regione r -esima; N_r è il numero di famiglie nella regione r -esima; γ è il

coefficiente di variazione desiderato delle stime; n_r^* è l'ampiezza del campione determinata dalle risorse disponibili.

Anziché valutare il coefficiente di variazione, C_{kr} , per ogni classe di spesa, in ogni regione è stato calcolato un solo coefficiente di variazione usando i dati dell'indagine sul consumo dell'Istat nel 1989: quello relativo alla spesa totale. Si è valutato che un errore relativo desiderato del 23%, che seppure elevato, produce una dimensione totale campionaria per periodo di registrazione delle spese pari a 1143 famiglie (si veda la tavola 4), che è quasi uguale a quella dell'indagine attuale in cui la dimensione prevista del campione è di 1100 USS.

Le valutazioni della dimensione dei campioni regionali mostrano che, se si tiene conto della stima della varianza della spesa totale, si dovrebbe sovracampionare nelle regioni meno popolate e con pochi Comuni grandi. Le dimensioni calcolate in base ai coefficienti di variazione della spesa totale mensile variano, sì, da regione a regione, ma hanno un ordine di grandezza quasi confrontabile tra loro; così, in regioni poco popolate, come la Valle d'Aosta, la Basilicata, l'Abruzzo, si ottiene un numero di famiglie da intervistare più alto rispetto a altre regioni. D'altronde, le dimensioni dei campioni regionali calcolati con l'allocatione proporzionale, n_{pr} , mostrano il difetto opposto: per le regioni meno popolate risultano notevolmente piccole. L'allocatione ottimale di Neyman (Cochran 1977), vincolata a un totale prefissato, non ha prodotto risultati diversi dall'allocatione proporzionale che perciò non sono stati riportati nella tavola 4; infatti, come si può dedurre dalla tavola A4 (nell'appendice A), la "variabilità" degli scarti quadratici medi della spesa totale da regione a regione è meno rilevante della variabilità del numero di famiglie. Se si considerano, invece, solo gli scarti quadratici medi come pesi per determinare l'allocatione tra le regioni della dimensione totale del campione, allora le dimensioni dei campioni regionali risultano circa uguali a quelli ottenuti in base alla precisione desiderata. Tali risultati sono coerenti con quanto già detto. Certamente, non è sensato che in Lombardia si campionino meno famiglie della Basilicata e non ha senso neanche campionare negli strati NAR una piccola quantità di USS perché si ottiene solo un aumento di variabilità.

I commenti sulle USS terminano qui. In seguito, si lavorerà solo sulle UPS che "trascinano" le USS corrispondenti. Rimane aperto, quindi, il

problema inerente a come fissare la numerosità delle USS per regione e a come allocarla tra le UPS della regione stessa. Si preferisce, in questa sede, affrontare esclusivamente la proposta di riduzione delle UPS.

Tavola 4 — *Popolazione residente (al 31/12/1989); numero di famiglie residenti (al 31/12/1988); coefficiente di variazione della spesa totale annua; dimensioni campionarie con errore relativo desiderato del 23%, proporzionali al numero di famiglie, medie per decade relative al 1989, e città con numero di abitanti residenti maggiore di 299.999, per regione.*

Regione	Abitanti	Famiglie	C_r	n_r	n_{pr}	\bar{n}	Città
Piemonte	4.357.559	1.757.792	0,815	48	98	64	TO
Valle d'Aosta	115.270	46.786	0,984	70	3	25	—
Lombardia	8.911.995	3.343.253	0,889	57	186	90	MI
Trentino-Alto-Adige	886.679	316.119	0,851	53	18	63	—
Veneto	4.385.023	1.525.734	0,816	48	85	48	VE
Friuli-Venezia-Giulia	1.202.877	484.682	0,892	58	27	29	—
Liguria	1.727.212	769.527	0,816	48	43	44	GE
Emilia-Romagna	3.921.597	1.503.938	0,830	50	84	56	BO
Toscana	3.560.582	1.323.431	0,860	54	74	64	FI
Umbria	820.316	299.989	0,784	45	17	27	—
Marche	1.430.726	496.715	0,795	46	28	42	—
Lazio	5.170.672	1.837.111	0,797	46	102	71	RM
Abruzzo	1.266.448	438.103	0,918	61	24	29	—
Molise	335.348	118.563	0,808	47	7	27	—
Campania	5.808.705	1.787.257	0,750	41	100	60	NA
Puglia	4.069.359	1.283.542	1,077	84	72	53	BA
Basilicata	623.175	207.702	1,040	79	12	26	—
Calabria	2.152.539	690.110	1,078	84	39	33	—
Sicilia	5.172.785	1.709.407	0,892	58	95	59	2(*)
Sardegna	1.657.562	527.410	0,955	66	29	29	—
Totale	57.576.429	20.467.171	0,888	1143	1143	939	11

(*) Palermo e Catania

4. Gli stimatori

Le modifiche proposte non influenzano le espressioni degli stimatori delle caratteristiche oggetto di studio. Infatti, sia Y_{thij} la spesa totale della j -esima famiglia nell' i -esimo Comune dell' h -esimo strato al tempo t . La spesa totale al tempo t , Y_t , è data da

$$Y_t = \sum_{h=1}^H \sum_{i=1}^{I_h} \sum_{j=1}^{N_{thi}} Y_{thij} \quad (1)$$

dove H indica il numero di strati, I_h il numero di Comuni nello strato h , N_{thi} il numero di famiglie nell' i -esimo Comune dello strato h al tempo t (si è utilizzato l'indice t per non appesantire la simbologia con un'ulteriore specificazione in decade e mese). Per determinare una stima della spesa si può ricorrere, per ogni strato, a uno stimatore del tipo:

$$\hat{Y}_t = \sum_{h=1}^H \frac{N_{th}}{n_{th}} \sum_{i=1}^{I_h} \left(\frac{N_{thi}}{n_{thi}} \sum_{j=1}^{n_{thi}} y_{thij} \right) \quad (2)$$

dove y_{thij} è il dato campionario della spesa, i_h il numero di Comuni nel campione nello strato h ; N_{th} , n_{th} , e n_{thi} indicano il numero di famiglie al tempo t nello strato h nella popolazione, nel campione, e nell' i -esimo Comune del campione, rispettivamente. La varianza dello stimatore è

$$V[\hat{Y}_t] = \sum_{h=1}^H \left(\frac{I_h}{i_h} \sum_{i=1}^{i_h} N_{thi}^2 \frac{s_{2thi}^2}{n_{thi}} (1 - f_{2thi}) \right) + \sum_{i=1}^{i_h} \left(\frac{I_h^2}{i_h} \sum_{i=1}^{i_h} N_{thi}^2 \frac{(\bar{y}_{thi} - \hat{Y}_{th})^2}{i_h - 1} (1 - f_{1th}) \right) \quad (3)$$

dove s_{2thi}^2 è la varianza campionaria della spesa, f_{2thi} la frazione di campionamento di famiglie nel Comune i -esimo, f_{1th} la frazione di campionamento dei Comuni, e \hat{Y}_{th} la spesa media nello strato h al tempo t così definita

$$\hat{Y}_{th} = \frac{\sum_{i=1}^{i_h} N_{thi} \bar{y}_{thi}}{\sum_{i=1}^{i_h} N_{thi}} \quad (4)$$

La spesa pro-capite, $\hat{Y}_{pc,t}$, si può ottenere dalla spesa totale al tempo t dividendola per la popolazione al tempo t , P_t . La varianza della spesa pro-

capite, $V[\hat{Y}_{pc,t}]$, allora avrà la stessa espressione di $V[\hat{Y}_t]$ della (3), moltiplicata per il quadrato dell'inverso della popolazione al tempo t , $(1/P_t)^2$. Alternativamente, si potrebbe usare la popolazione stessa come "peso" per determinare la stima del consumo totale e si avrebbe:

$$\hat{Y}_{pc,t} = \frac{1}{P_t} \left[\sum_{h=1}^H \left[\frac{P_{th}}{P_t} \sum_{i=1}^{i_h} \left(\frac{P_{thi}}{P_{th}} \sum_{j=1}^{n_{thi}} y_{thij} \right) \right] \right] \quad (5)$$

dove P_{th} e P_{thi} indicano la popolazione residente al tempo t nello strato h , e al tempo t nello strato h nel Comune i , rispettivamente; mentre p_{th} e p_{thi} indicano la popolazione nel campione al tempo t nello strato h , al tempo t nello strato h nel Comune i , rispettivamente. La varianza dello stimatore è

$$V[\hat{Y}_{pc,t}] = \left(\frac{1}{P_t} \right)^2 \sum_{h=1}^H \left(\frac{I_{th}}{i_{th}} \sum_{i=1}^{i_h} P_{thi}^2 \frac{S_{2thi}^2}{n_{thi}} (1 - f_{2thi}) \right) + \left(\frac{1}{P_t} \right)^2 \sum_{h=1}^H \left(\frac{I_{th}^2}{i_{th}} P_{thi}^2 \frac{(\bar{y}_{thi} - \hat{Y}_{th})^2}{i_{th} - 1} (1 - f_{ith}) \right). \quad (6)$$

Una diminuzione del numero di unità di primo stadio negli strati, i_{th} , produce una maggiore instabilità nella stima della variabilità dei caratteri tra le unità. Il fenomeno ha particolare rilievo per la classe 1, dove può risultare conveniente selezionare le unità con probabilità proporzionale alla loro dimensione. Gli stimatori non si modificano radicalmente perché il piano di campionamento conserva la stessa "struttura" dell'indagine attuale, le rotazioni dei gruppi avvengono in periodi multipli di un trimestre e riguardano solo le UPS.

Quanto precede rappresenta una semplificazione delle espressioni perché il tempo t andrebbe esplicitato meglio in decenni, mesi, e trimestri. Inoltre, le unità sono state trattate come se fossero omogenee per strato. Un miglioramento ulteriore delle stime si ottiene con la post-stratificazione sull'ampiezza della famiglia (Falorsi *et al.* 1992).

5. Alcune simulazioni basate sui dati dell'indagine del 1989

Si è eseguita una simulazione per accertare come le stime e i corrispondenti scarti quadratici medi si modificano al variare del numero di UPS preso in considerazione. I dati dell'indagine Istat del 1989 sono quindi stati usati come una popolazione da cui sono state estratte le UPS secondo la stratificazione proposta. La simulazione quindi non ha riguardato la scelta delle USS, che sono state sempre associate alla propria UPS, se selezionata. Poiché i dati disponibili della popolazione non erano disaggregati nel tempo, gli stimatori riportati nel paragrafo 4 sono stati lievemente modificati (paragrafo 5.1); tali dati non erano disaggregati neanche per ampiezza della famiglia sicché non si è potuta applicare una post-stratificazione per migliorare le stime (Falorsi *et al.* 1992). Le simulazioni hanno riguardato, quindi, esclusivamente le UPS e sono state eseguite per conseguire due obiettivi diversi: 1) studiare il comportamento delle stime e dei relativi scarti quadratici medi all'aumentare delle UPS incluse nel campione secondo diversi criteri di inclusione (paragrafo 5.2), 2) calcolare le stime e gli scarti quadratici medi inerenti al piano di campionamento proposto per valutare la stabilità delle stime (paragrafo 5.3).

5.1. Stimatori utilizzati nella simulazione

Nel paragrafo 4 si sono presentati gli stimatori nell'accezione più generale. Per calcolare le stime secondo la modifica al piano di campionamento occorrono alcune informazioni relative alla popolazione oggetto di indagine. I dati disponibili, però, riguardano solo la popolazione relativa alle UPS incluse nel campione e il numero totale di famiglie per regione sia delle USS, sia della popolazione totale; ma non è nota, per esempio, la popolazione stratificata secondo l'ampiezza delle USS. Se si tiene conto della struttura dell'indagine Istat sul consumo e dei dati disponibili, si possono "correggere" gli stimatori visti in precedenza al fine di calcolare, nella fase di simulazione, le stime e gli scarti quadratici medi della spesa per il consumo totale, per il consumo di beni alimentari e non alimentari. Gli indici e le grandezze utilizzate sono state riportate nella tavola 5.

Tavola 5 — Indici e grandezze utilizzate nelle espressioni degli stimatori.

d	decade	$d=1,2,3$
m	mese	$m=1,2,3$
t	trimestre	$t=1,2,3,4$
h	strato	$h=1,...,H$
i	Comune (UPS)	$i=1,...,I_h$
j	famiglia (USS)	$j=1,...,J_{ih}$

Si noti che nella stratificazione attuale i_{tmh} è uguale a 1 per ogni t, m, d .

La spesa per il consumo totale nel campione al trimestre t , al mese m , nello strato h , dell' i -esima UPS, y_{tmhi} , sarà uguale alla somma delle spese osservate nelle USS selezionate nelle 3 decadi, y_{tmhdij} , e successivamente intervistate

$$y_{tmhi} = \sum_{d=1}^3 \sum_{j=1}^{n_{tmhdij}} y_{tmhdij},$$

dove l'indice j denota la j -esima famiglia e n_{tmhdij} è il numero di famiglie intervistate al trimestre t (mese m e decade d), nello strato h e nell' i -esima UPS. Inoltre, l'ammontare della popolazione nel campione al trimestre t , al mese m , nello strato h , dell' i -esima UPS, p_{tmhi} , sarà uguale a

$$p_{tmhi} = \sum_{d=1}^3 \sum_{j=1}^{n_{tmhdij}} p_{tmhdij},$$

dove p_{tmhdij} è l'ampiezza della j -esima famiglia nell' i -esima UPS dello strato h al trimestre t (mese m e decade d). Pertanto, se si indicano con P i valori dell'entità della popolazione (numero di abitanti), il consumo pro-capite nel campione e la spesa stimata per il consumo totale al trimestre t , al mese m , nello strato h , dell' i -esima UPS, saranno uguali, rispettivamente, a

$$\hat{y}_{pc,tmhi} = \frac{y_{tmhi}}{p_{tmhi}} \quad \text{e} \quad \hat{Y}_{pc,tmhi} = P_{tmhi} \frac{y_{tmhi}}{p_{tmhi}}.$$

Se si indica con $P_{tmh}^c = \sum_{i=1}^{i_{tmh}} P_{tmhi}$ la popolazione totale delle UPS del campione, allora la somma delle spese stimate per il consumo totale nelle

UPS incluse nel campione divisa per P_{tmh}^c dà ancora la spesa stimata per il consumo pro-capite. La spesa per il consumo totale stimata al trimestre t , al mese m , nello strato h , quindi, sarà data da

$$\hat{Y}_{tmh} = \frac{P_{tmh}}{P_{tmh}^c} \sum_{i=1}^{i_{tmh}} P_{tmhi} \frac{y_{tmhi}}{p_{tmhi}} \quad (7)$$

Lo stimatore della spesa per il consumo totale trimestrale si ottiene sommando per ogni strato e per ogni mese l'espressione precedente:

$$\hat{Y}_t = \sum_{m=1}^3 \sum_{h=1}^H \frac{P_{tmh}}{P_{tmh}^c} \sum_{i=1}^{i_{tmh}} P_{tmhi} \frac{y_{tmhi}}{p_{tmhi}}. \quad (8)$$

La varianza dello stimatore ottenuto si semplifica molto perché per ogni mese vi è una sola UPS per strato e risulta uguale alla seguente:

$$V(\hat{Y}_t) = \sum_{m=1}^3 \sum_{h=1}^H \left(\frac{P_{tmh}}{P_{tmh}^c} \right)^2 \sum_{i=1}^{i_{tmh}} \left(\frac{P_{tmhi}}{p_{tmhi}} \right)^2 n_{tmhi} s_{tmhi}^2 (1 - f_{2tmhi}).$$

Tale espressione si ottiene tenendo conto che y_{tmhi} può essere riscritta come prodotto del numero di famiglie nel campione, relativo all' i -esima UPS dello strato h nel mese m del trimestre t , per la media campionaria corrispondente: $n_{tmhi} \bar{y}_{tmhi}$. Si noti che non figura il secondo termine della varianza riportata nel paragrafo precedente, che indicava la variabilità tra le UPS dello strato, perché per ogni strato e per ogni mese si seleziona una sola UPS.

In questo procedimento per stimare il consumo totale si è partiti dalla stima del consumo pro-capite. Si è così impostata la stima rispetto alla popolazione (numero di abitanti) e non rispetto alle USS perché si è voluto tenere conto implicitamente dell'ampiezza della famiglia. Diversamente, si ottengono formule analoghe alle precedenti sostituendo le lettere P o p , che indicano la popolazione con le lettere N o n , rispettivamente, che indicano il numero di famiglie nel paese (N) e nel campione (n).

5.2. *L'importanza dell'ordine di inclusione delle unità di primo stadio nel campione*

Una prima simulazione ha estratto in modo sequenziale dal campione corrente dell'Istat una successione di campioni di ampiezza crescente (fino a raggiungere le dimensioni del campione attuale) selezionando le UPS per regione con criteri diversi. Da ciascun campione estratto sono state calcolate le stime delle voci di spesa considerate. L'inserimento delle UPS è stato effettuato separatamente per le singole regioni perché il livello di precisione della spesa doveva essere riferito specificamente al livello regionale. Attualmente in ogni strato le UPS di tipo NAR sono tre in un anno e sono state considerate come replicazioni della stessa unità. Si sono seguiti tre criteri diversi di inclusione delle UPS.

5.2.1. *Inclusione delle UPS in ordine decrescente*

Il primo criterio consiste nella selezione successiva delle UPS ordinate in senso decrescente rispetto alla loro ampiezza (il numero di abitanti), distinguendo tra AR e NAR. In altri termini, se i è il numero di UPS selezionate per regione, l'inclusione avviene nel seguente modo.

Per $i=1$ in ogni regione si seleziona una sola UPS (la più popolata); si ottiene così un sotto-campione dal quale si determinano le stime e gli scarti quadratici medi corrispondenti delle voci di spesa considerate. Per $i=2$ si aggiunge al sotto-campione precedente una UPS selezionata in ogni regione tra quelle appartenenti alle attuali AR (se le UPS di tipo AR sono "esaurite", come nel caso della Valle d'Aosta, la regione rimane con una sola UPS); si ottiene così un nuovo sotto-campione dal quale si determinano le stime e gli scarti quadratici medi corrispondenti. Si procede così fino a $i=16$, ossia fino a quando tutte le UPS di tipo AR non sono state inserite; infatti, la Campania conta 16 UPS di tipo AR. Con lo stesso procedimento si includono le UPS di tipo NAR fino a $i=29$, ossia fino a quando tutte le UPS di tipo NAR non sono state incluse; infatti, la Lombardia conta 13 UPS di tipo NAR. La percentuale di UPS incluse per ogni i è riportata nel grafico 8 dell'appendice B e si può evincere anche dalla tavola 6 (esposta nel successivo paragrafo 5.3) in cui è riportata la suddivisione in classi delle UPS per regione e per tipo AR-NAR.

5.2.2. *Inclusione per prima della UPS più grande, e le altre in ordine crescente*

Il secondo criterio consiste nell'inclusione successiva delle UPS ordinate in senso crescente rispetto alla loro ampiezza (il numero di abitanti), distinguendo tra AR e NAR, eccetto la prima UPS inclusa che riguarda le unità più popolate e corrispondenti alla classe 4 della stratificazione proposta nel paragrafo 3 (per la Sicilia si include solo Palermo). In altri termini, se i è il numero di UPS selezionate per regione, l'inclusione avviene nel seguente modo.

Per $i=1$ in ogni regione si seleziona una sola UPS (la più popolata); si ottiene così un sotto-campione dal quale si determinano le stime e gli scarti quadratici medi corrispondenti delle voci di spesa considerate. Per $i=2$ si aggiunge al sotto-campione precedente una UPS selezionata in ogni regione tra quelle appartenenti alle attuali NAR con la dimensione più piccola; si ottiene così un nuovo sotto-campione dal quale si determinano le stime e gli scarti quadratici medi corrispondenti. Per $i=3$ si aggiunge al sotto-campione precedente una UPS selezionata in ogni regione tra quelle con dimensione più piccola tra le rimanenti NAR (quando le unità NAR sono "esaurite", la regione mantiene sempre lo stesso numero di UPS); si ottiene così un nuovo sotto-campione dal quale si determinano le stime e gli scarti quadratici medi corrispondenti. Si procede così fino a $i=14$, ossia fino a quando tutte le UPS di tipo NAR non sono state inserite; infatti, la Lombardia conta 13 UPS di tipo NAR. Con lo stesso procedimento si includono, a partire dalle più piccole, le UPS di tipo AR fino a $i=29$, ossia fino a quando tutte le UPS di tipo AR non sono state incluse; infatti, la Campania conta altre 15 UPS di tipo AR (l'unità AR più popolata è stata inclusa per $i=1$).

5.2.3. *Inclusione delle UPS in ordine crescente*

Il terzo criterio esaminato consiste nella selezione successiva delle UPS ordinate in senso decrescente rispetto alla loro ampiezza (il numero di abitanti), distinguendo tra AR e NAR. In altri termini, se i è il numero di UPS selezionate per regione, l'inclusione avviene nel seguente modo.

risultando in genere più vicina al minimo; il campo di variazione delle medie è, rispetto all'anno e alle stime ottenute dal campione attuale, sempre dell'ordine del 4% per la spesa totale, dell'1% per la spesa alimentare, del 5% per la spesa non alimentare.

Tavola 10 — *Stime e scarti quadratici medi relativi al campione totale Istat 1989 e ai 1000 campioni generati secondo lo schema proposto tavola 3, media (minima e massima) e efficienza (minima, media e massima) per il consumo totale, per il consumo alimentare e per il consumo non alimentare.*

	I trimestre	II trimestre	III trimestre	IV trimestre	Anno
Consumo t.					
UPS totali	44,84727200	47,44197441	48,91290488	53,84206258	195,04421387
d.s. UPS totali	0,58587106	0,71295070	0,73392359	0,67124322	1,35674423
med. simulaz.	44,31704594	47,62665906	48,99064480	53,66964119	194,60399099
d.s. simulaz.	0,50768065	0,53536460	0,42435789	0,50398472	1,12099774
media min	42,92833530	46,03723964	47,74790354	52,16752744	191,53993837
media max	45,82443302	49,36337081	50,55701545	55,21286151	198,12905149
I.S.S. min	1,1931	1,1496	1,2488	1,2878	1,3151
I.S.S. medio	1,5118	1,4790	1,5612	1,5654	1,5552
I.S.S. max	2,0336	2,2807	2,1275	1,9662	1,7999
C. Alimentare					
UPS totali	10,81383213	11,08565781	11,18700780	12,18412421	45,27062195
d.s. UPS totali	0,10210434	0,09944523	0,10670001	0,10977737	0,20916647
med. simulaz.	10,79621774	11,20333763	11,25827365	12,21604820	45,47387722
d.s. simulaz.	0,09081839	0,10270224	0,09464256	0,12459238	0,24166926
media min	10,51964115	10,92470529	10,95234740	11,83823667	44,77068283
media max	11,07040257	11,50870478	11,47297116	12,49148178	46,09298393
I.S.S. min	1,3227	1,3152	1,3322	1,3934	1,4720
I.S.S. medio	1,6507	1,6334	1,5833	1,6695	1,6367
I.S.S. max	2,3140	1,9981	2,0829	2,4361	1,9179
C. non alim.					
UPS totali	34,03343987	36,35631660	37,72589708	41,65793837	149,77359192
d.s. UPS totali	0,54414126	0,67395418	0,69258064	0,62395539	1,27251413
med. simulaz.	33,52082820	36,42332143	37,73237115	41,45359299	149,13011377
d.s. simulaz.	0,47098172	0,47676251	0,38313921	0,43291824	0,99730139
media min	32,21329460	34,97420545	36,56286640	40,17969017	146,42055946
media max	34,93109819	37,94925442	39,20571500	42,78281917	152,28948562
I.S.S. min	1,1441	1,1118	1,2335	1,2809	1,2958
I.S.S. medio	1,4901	1,5586	1,5654	1,5544	1,5489
I.S.S. max	2,0167	2,3037	2,1748	1,9186	1,7937

I risultati conseguiti con queste simulazioni confermano l'elevata variabilità della spesa non alimentare che rappresenta la componente più sensibile alle caratteristiche socio-economiche delle USS e, quindi, più

dispersa. La scarsa concentrazione della spesa non alimentare richiede necessariamente una maggiore attenzione nell'accertamento perché essa è composta anche da spese che sono di difficile rilevazione sia per cause "oggettive" connesse alla rarità degli eventi o alla complessità del fenomeno (Drudi 1992), sia per cause "soggettive" connesse alla reticenza o disattenzione dell'intervistato.

6. Conclusioni

In questo lavoro ci siamo proposti di esaminare le conseguenze di alcune variazioni del piano di campionamento volte alla diminuzione delle unità di primo e secondo stadio al fine di aumentare le risorse e favorire le strategie per controllare la qualità dei dati. I risultati ottenuti con lo schema di campionamento illustrato sono soddisfacenti, anche se tale schema rappresenta comunque solo uno spunto per discussioni più ampie e approfondite. Infatti, i Comuni NAR potrebbero essere, per esempio, ulteriormente stratificati sull'ampiezza demografica (soprattutto se la si considera come una *proxy* del consumo). Altri fattori determinanti per la stima dei consumi, e quindi per il piano di campionamento, sono il reddito, l'età e il sesso dei componenti della famiglia, la condizione professionale del capofamiglia e il suo livello di istruzione; ma allo stato attuale sembra difficile poter stratificare su tali caratteri perché lo sforzo organizzativo necessario per avere questi dati aggiornati in ogni area geografica o Comune è notevole.

La revisione dell'indagine sui consumi e sui bilanci delle famiglie richiede anche la valutazione di metodi alternativi all'attuale indagine ripetuta perché offrono vantaggi sia nell'acquisizione dei dati, sia nella formulazione dei modelli, come avviene in alcuni paesi: Stati Uniti (Survey Research Center 1984) e Germania (Hanefeld 1984). Tra i vari metodi, l'indagine panel costituisce un'alternativa interessante (Duncan e Kalton 1987; Kish 1989; Kasprzyk *et al.* 1989), anche quando usa un campione limitato nel tempo a due o tre anni (Lalla 1992). L'indagine *split panel* (Kish 1983, 1986), che utilizza un campione con due componenti distinte e "indipendenti": una panel e l'altra rotata o ripetuta, combina i vantaggi dei due metodi; tramite la componente panel si stimano la variazione netta e i cambiamenti individuali, si cumulano i dati

Grafico 11 – Stime della spesa per il consumo di beni alimentari al variare del numero di UPS nel campione in base ai tre criteri di inclusione

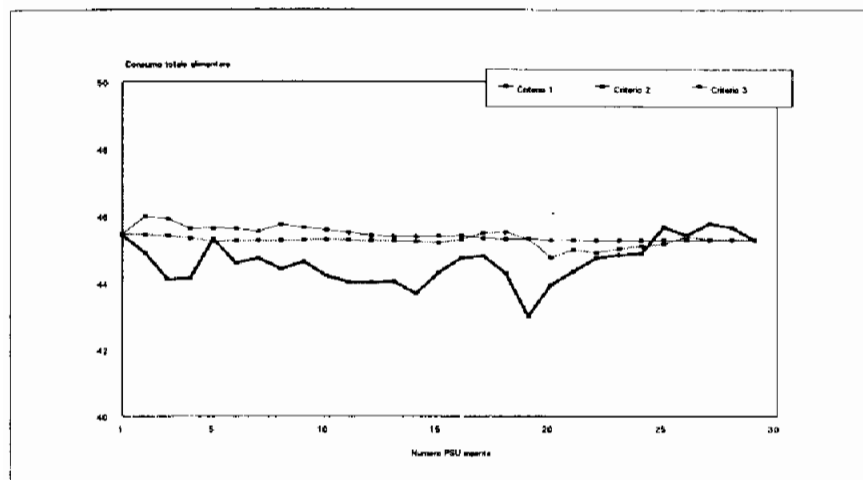


Grafico 12 – Deviazioni standard delle stime della spesa per il consumo di beni alimentari al variare del numero di UPS inserite nel campione in base ai tre criteri di inclusione

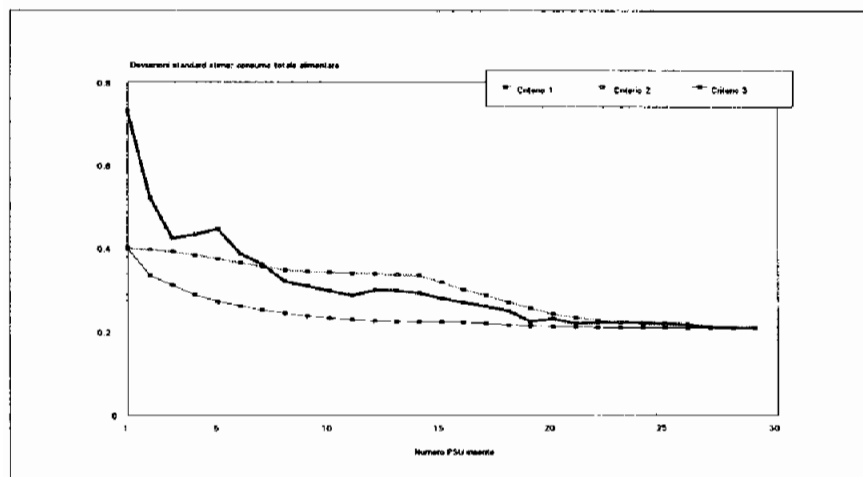


Grafico 13 – Stime della spesa per il consumo di beni non alimentari al variare del numero di UPS nel campione in base ai tre criteri di inclusione

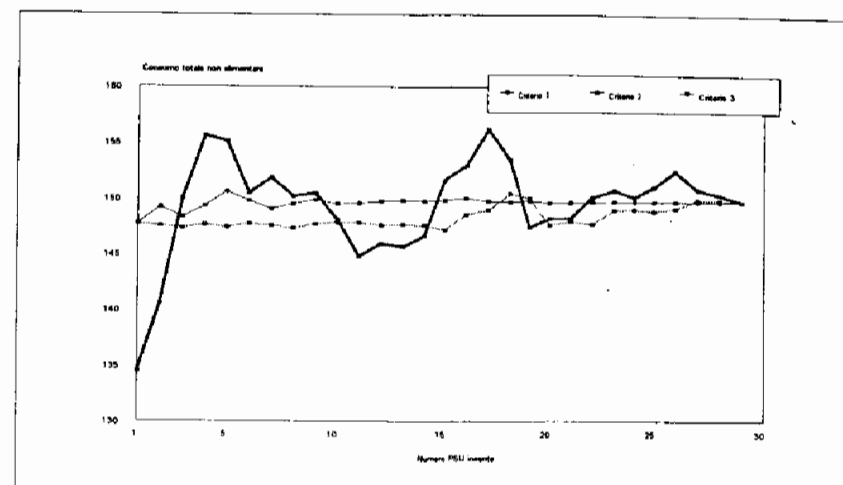
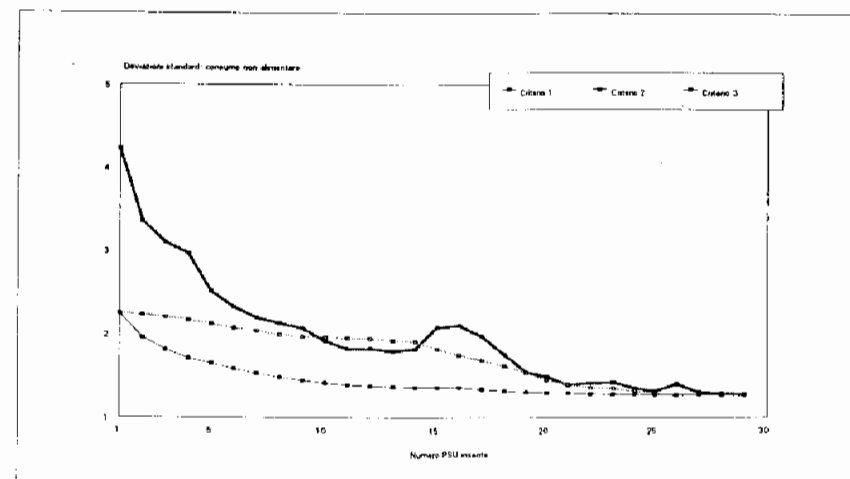


Grafico 14 – Deviazioni standard delle stime della spesa per il consumo di beni non alimentari al variare del numero di UPS inserite nel campione in base ai tre criteri di inclusione



attribuite a tre gruppi *a*, *b*, *c*, definiti in precedenza; quindi, entrano nel campione secondo lo schema previsto: il gruppo *a* rimane nel campione i primi sei mesi, e il gruppo *b* entra nei sei mesi successivi (il gruppo *c* viene escluso). Le UPS della classe I vengono scelte casualmente secondo il numero previsto (due per strato) e vengono suddivise in due gruppi: il primo rimane nel campione per i primi sei mesi e il secondo entra nei sei mesi successivi (in realtà, ogni gruppo "semestrale" è formato da tre UPS dei NAR attuali).

Tavola 6 — *UPS dell'indagine Istat attuale suddivise per classi di ampiezza secondo lo schema proposto.*

Regione	UPS e x AR	UPS e x NAR
Piemonte	4 3 2 2 2 2 2	1 1 1 1 1 1 1 1 1
Valle d'Aosta	4	1 1 1 1
Lombardia	4 3 3 3 2 2 2 2 2 2 2 2 1	1 1 1 1 1 1 1 1 1 1 1 1
Trentino-Alto-Adige	4 3	1 1 1 1 1 1 1 1
Veneto	4 3 3 3 2 2 2 1	1 1 1 1 1 1 1 1
Friuli-Venezia-Giulia	4 3 2 1	1 1 1
Liguria	4 3 2 2 1	1 1 1 1 1
Emilia-Romagna	4 3 3 3 3 3 3 3 2 2 2	1 1 1 1 1 1 1 1
Toscana	4 3 3 3 2 2 2 2 2 2 2 2	1 1 1 1 1 1 1 1
Umbria	4 3 2	1 1 1 1
Marche	4 2 2 2 1	1 1 1 1 1
Lazio	4 3 2 2 2 2 1 1	1 1 1 1 1 1
Abruzzo	4 2 2 2	1 1 1 1 1
Molise	4 1	1 1 1 1 1
Campania	4 3 3 2 2 2 2 2 2 2 2 2 2 2	1 1 1 1 1 1 1 1 1 1
Puglia	4 3 3 3 2 2 2 2 2 2 2 2	1 1 1 1 1 1 1 1
Basilicata	4 2	1 1 1 1
Calabria	4 3 3 2 2	1 1 1 1 1 1
Sicilia	4 4 3 3 2 2 2 2 2 2 1	1 1 1 1 1 1 1 1
Sardegna	4 3 2 1 1	1 1 1 1 1

Ottenuto il campione, si determinano le stime. Quindi, il ciclo si ripete: si estrae un altro campione, si determinano altre stime; e così via di seguito. Il procedimento è stato ripetuto 1000 volte. Nella tavola 7 compaiono alcuni indicatori riassuntivi delle simulazioni effettuate: il valore medio sui 1000 casi della percentuale di UPS (e USS) che entra in ciascun campione, il minimo e il massimo di tale percentuale.

Nella tavola 8 si sono riportate, per ogni tipo di spesa, le stime del totale e i relativi scarti quadratici medi, rispettivamente, per il campione Istat e per i 1000 campioni generati, il minimo e il massimo di tali medie. Le stime della spesa per le tre variabili considerate sono state calcolate con la media aritmetica semplice delle 1000 stime derivate dagli altrettanto campioni di UPS selezionate casualmente, mentre i corrispondenti scarti quadratici medi sono stati ottenuti con la formula usuale (Efron e Tibshirani 1993, pp. 39-56). I valori ottenuti sono risultati lievemente inferiori (circa il 10%) degli scarti quadratici medi ottenuti dal campione totale dell'indagine del 1989 e inferiori (circa il 30%) degli scarti quadratici medi ottenuti con la media aritmetica semplice dei 1000 scarti quadratici medi derivati dai 1000 campioni di UPS selezionate casualmente (dati non riportati nelle tavole).

Tavola 7 — *Percentuali di inclusioni di UPS e USS (medie, minime e massime) risultate da 1000 campioni, generati dal campione Istat del 1989, simili a quello proposto nella tavola 3, ma con una sola UPS per strato ottenuto stratificando la classe I per l'altezza s.l.m. delle UPS.*

	I trimestre	II trimestre	III trimestre	IV trimestre	Anno
m. % UPS in	43,54	43,80	43,72	43,63	43,67
% min UPS in	33,76	33,92	34,00	34,63	39,09
% max UPS in	51,87	51,75	52,80	52,35	47,84
m. % USS in	49,54	50,12	49,54	49,64	49,71
% min USS in	39,86	40,70	40,33	40,88	46,12
% max USS in	58,33	58,25	57,59	57,78	53,66

Per valutare l'effetto che il piano di campionamento modificato ha sulle stime, rispetto al piano di campionamento attuale, si è calcolato il rapporto tra la varianza delle stime ottenute da un campione generato da una simulazione e la varianza del piano di campionamento attuale. Si denoterà tale rapporto, che determina l'efficienza relativa delle due stime, come *effetto del disegno proposto*, anche se non viene calcolato nel modo tradizionale in cui il denominatore si riferisce alla varianza dello stimatore in un campionamento casuale semplice di pari dimensione (Kish 1965; Cicchitelli, Herzel e Montanari 1992); inoltre, la varianza del numeratore "dipende" dalla varianza del campione dato dall'indagine del 1989. Nella tavola 8 l'effetto del disegno proposto si è indicato con l'acronimo I.S.S. (Indice di Stabilità delle Stime) perché esprime la

variabilità delle stime relativamente al campione (dell'indagine del 1989) utilizzato nelle simulazioni.

Tavola 8 — *Stime e scarti quadratici medi relativi al campione totale Istat 1989 e ai 1000 campioni generati secondo lo schema proposto tavola 3 (con una sola UPS per strato ottenuto stratificando la classe 1 per l'altezza s.l.m. delle UPS), media (minima e massima) e effetto del disegno (minimo, medio e massimo) per il consumo totale, per il consumo alimentare e per il consumo non alimentare.*

	I trimestre	II trimestre	III trimestre	IV trimestre	Anno
Consumo t.					
UPS totali	44,84727200	47,44197441	48,91290488	53,84206258	195,04421387
d.s. UPS totali	0,58587106	0,71295070	0,73392359	0,67124322	1,35674423
med. simulaz.	44,27920390	47,60319377	49,01997913	53,72843738	194,63081419
d.s. simulaz.	0,52277496	0,53304814	0,43469102	0,47267094	1,11937937
media min	42,74593295	45,98733222	47,58205381	52,26319638	191,23988192
media max	45,83844125	49,45916267	50,53586725	55,31344634	198,15184819
I.S.S. min	1,2061	1,2420	1,3432	1,3446	1,4056
I.S.S. medio	1,5560	1,6165	1,6172	1,6369	1,6121
I.S.S. max	2,1484	2,2093	2,0582	2,1577	1,8663
C. alimentare					
UPS totali	10,81383213	11,08565781	11,18700780	12,18412421	45,27062195
d.s. UPS totali	0,10210434	0,09944523	0,10670001	0,10977737	0,20916647
med. simulaz.	10,80241167	11,21589177	11,27039165	12,23163995	45,52033503
d.s. simulaz.	0,09562602	0,10431474	0,09581812	0,11988603	0,23703293
media min	10,47211243	10,91660643	10,95990274	11,83122281	44,75210410
media max	11,11406002	11,52694249	11,52833552	12,57084336	46,16940361
I.S.S. min	1,4185	1,3830	1,3932	1,4093	1,4986
I.S.S. medio	1,7242	1,8877	1,6424	1,7476	1,7062
I.S.S. max	2,2606	2,2123	1,9848	2,3565	1,9591
C. non alim.					
UPS totali	34,03343987	36,35631660	37,72589708	41,65793837	149,77359192
d.s. UPS totali	0,54414126	0,67395418	0,69258064	0,62395539	1,27251413
med. simulaz.	33,47679223	36,38730201	37,74958748	41,49679743	149,11047916
d.s. simulaz.	0,48563176	0,47873293	0,39732135	0,41724077	1,01423782
media min	32,03539410	34,82438596	36,61903524	40,28705114	145,92236702
media max	35,00625815	37,96328334	39,18768488	42,85552040	152,22112002
I.S.S. min	1,1664	1,2283	1,3367	1,3322	1,3927
I.S.S. medio	1,5297	1,6112	1,6210	1,6257	1,6043
I.S.S. max	2,1229	2,2411	2,0759	2,1316	1,8664

Il risultato ottenuto è soddisfacente perché, nei casi peggiori corrispondenti al massimo dell'effetto del disegno, si può vedere che l'efficienza relativa delle stime è doppia rispetto al campione attuale dell'indagine sui consumi dell'Istat; il valore medio mostra una lieve

asimmetria perché tende a essere più vicino al minimo; il campo di variazione delle stime ottenute con la simulazione è, rispetto all'anno e alla stima derivante dal campione attuale, dell'ordine del 4% per la spesa totale, dell'1% per la spesa alimentare, e del 5% per la spesa non alimentare.

In modo analogo al precedente schema di simulazione si è esaminato il caso in cui nella classe 1 si sceglie il numero di UPS uguale a quello indicato nella tavola 3 (l'ottava colonna indicata con classe 1). Rispetto al procedimento sopra descritto, cambia solo il numero di UPS selezionate nella classe 1: le UPS vengono scelte casualmente in numero doppio di quello previsto dall'ottava colonna della tavola 3 e suddivise in due gruppi: il primo rimane nel campione per i primi sei mesi e il secondo entra nei sei mesi successivi (ogni UPS del gruppo "semestrale" è sempre formata da tre UPS dei NAR attuali). Il procedimento di simulazione è identico al precedente e i risultati sono riportati nella tavola 9, contenente le statistiche riassuntive relative all'inclusione delle UPS e USS, e nella tavola 10 (analoga alla tavola 8), contenente le stime delle spese considerate, gli scarti quadratici medi relativi, l'effetto del disegno medio, e il minimo e il massimo delle stime e dell'effetto del disegno corrispondenti ai 1000 campione generati.

Tavola 9 — *Percentuali di inclusioni di UPS e USS (medie, minime e massime) risultate da 1000 campioni, generati dal campione Istat del 1989, simili a quello proposto nella tavola 3.*

	I trimestre	II trimestre	III trimestre	IV trimestre	Anno
m. % UPS in	50,98	51,19	51,30	51,20	51,17
% min UPS in	39,37	40,79	41,75	41,97	47,62
% max UPS in	58,62	60,09	59,97	60,39	54,91
m. % USS in	57,11	57,63	57,37	57,45	57,39
% min USS in	47,02	47,89	48,76	49,06	55,00
% max USS in	64,71	64,84	66,16	65,29	60,23

Le grandezze riportate nella tavola 10 mostrano esiti sostanzialmente uguali al caso precedente e talvolta si ottengono lievi peggioramenti, nonostante l'aumento delle UPS considerate nel campione derivato con la simulazione dal campione attuale dell'indagine sui consumi dell'Istat. L'efficienza massima relativa al campione attuale rimane dell'ordine del doppio e l'efficienza media mostra ancora una lieve asimmetria

risultando in genere più vicina al minimo; il campo di variazione delle medie è, rispetto all'anno e alle stime ottenute dal campione attuale, sempre dell'ordine del 4% per la spesa totale, dell'1% per la spesa alimentare, del 5% per la spesa non alimentare.

Tavola 10 — *Stime e scarti quadratici medi relativi al campione totale Istat 1989 e ai 1000 campioni generati secondo lo schema proposto tavola 3, media (minima e massima) e efficienza (minima, media e massima) per il consumo totale, per il consumo alimentare e per il consumo non alimentare.*

	I trimestre	II trimestre	III trimestre	IV trimestre	Anno
Consumo t.					
UPS totali	44,84727200	47,44197441	48,91290488	53,84206258	195,04421387
d.s. UPS totali	0,58587106	0,71295070	0,73392359	0,67124322	1,35674423
med. simulaz.	44,31704594	47,62665906	48,99064480	53,66964119	194,60399099
d.s. simulaz.	0,50768065	0,53536460	0,42435789	0,50398472	1,12099774
media min	42,92833530	46,03723964	47,74790354	52,16752744	191,53993837
media max	45,82443302	49,36337081	50,55701545	55,21286151	198,12905149
I.S.S. min	1,1931	1,1496	1,2488	1,2878	1,3151
I.S.S. medio	1,5118	1,4790	1,5612	1,5654	1,5552
I.S.S. max	2,0336	2,2807	2,1275	1,9662	1,7999
C. Alimentare					
UPS totali	10,81383213	11,08565781	11,18700780	12,18412421	45,27062195
d.s. UPS totali	0,10210434	0,09944523	0,10670001	0,10977737	0,20916647
med. simulaz.	10,79621774	11,20333763	11,25827365	12,21604820	45,47387722
d.s. simulaz.	0,09081839	0,10270224	0,09464256	0,12459238	0,24166926
media min	10,51964115	10,92470529	10,95234740	11,83823667	44,77068283
media max	11,07040257	11,50870478	11,47297116	12,49148178	46,09298393
I.S.S. min	1,3227	1,3152	1,3322	1,3934	1,4720
I.S.S. medio	1,6507	1,6334	1,5833	1,6695	1,6367
I.S.S. max	2,3140	1,9981	2,0829	2,4361	1,9179
C. non alim.					
UPS totali	34,03343987	36,35631660	37,72589708	41,65793837	149,77359192
d.s. UPS totali	0,54414126	0,67395418	0,69258064	0,62395539	1,27251413
med. simulaz.	33,52082820	36,42332143	37,73237115	41,45359299	149,13011377
d.s. simulaz.	0,47098172	0,47676251	0,38313921	0,43291824	0,99730139
media min	32,21329460	34,97420545	36,56286640	40,17969017	146,42055946
media max	34,93109819	37,94925442	39,20571500	42,78281917	152,28948562
I.S.S. min	1,1441	1,1118	1,2335	1,2809	1,2958
I.S.S. medio	1,4901	1,5586	1,5654	1,5544	1,5489
I.S.S. max	2,0167	2,3037	2,1748	1,9186	1,7937

I risultati conseguiti con queste simulazioni confermano l'elevata variabilità della spesa non alimentare che rappresenta la componente più sensibile alle caratteristiche socio-economiche delle USS e, quindi, più

dispersa. La scarsa concentrazione della spesa non alimentare richiede necessariamente una maggiore attenzione nell'accertamento perché essa è composta anche da spese che sono di difficile rilevazione sia per cause "oggettive" connesse alla rarità degli eventi o alla complessità del fenomeno (Drudi 1992), sia per cause "soggettive" connesse alla reticenza o disattenzione dell'intervistato.

6. Conclusioni

In questo lavoro ci siamo proposti di esaminare le conseguenze di alcune variazioni del piano di campionamento volte alla diminuzione delle unità di primo e secondo stadio al fine di aumentare le risorse e favorire le strategie per controllare la qualità dei dati. I risultati ottenuti con lo schema di campionamento illustrato sono soddisfacenti, anche se tale schema rappresenta comunque solo uno spunto per discussioni più ampie e approfondite. Infatti, i Comuni NAR potrebbero essere, per esempio, ulteriormente stratificati sull'ampiezza demografica (soprattutto se la si considera come una *proxy* del consumo). Altri fattori determinanti per la stima dei consumi, e quindi per il piano di campionamento, sono il reddito, l'età e il sesso dei componenti della famiglia, la condizione professionale del capofamiglia e il suo livello di istruzione; ma allo stato attuale sembra difficile poter stratificare su tali caratteri perché lo sforzo organizzativo necessario per avere questi dati aggiornati in ogni area geografica o Comune è notevole.

La revisione dell'indagine sui consumi e sui bilanci delle famiglie richiede anche la valutazione di metodi alternativi all'attuale indagine ripetuta perché offrono vantaggi sia nell'acquisizione dei dati, sia nella formulazione dei modelli, come avviene in alcuni paesi: Stati Uniti (Survey Research Center 1984) e Germania (Hanefeld 1984). Tra i vari metodi, l'indagine panel costituisce un'alternativa interessante (Duncan e Kalton 1987; Kish 1989; Kasprzyk *et al.* 1989), anche quando usa un campione limitato nel tempo a due o tre anni (Lalla 1992). L'indagine *split panel* (Kish 1983, 1986), che utilizza un campione con due componenti distinte e "indipendenti": una panel e l'altra rotata o ripetuta, combina i vantaggi dei due metodi; tramite la componente panel si stimano la variazione netta e i cambiamenti individuali, si cumulano i dati

delle unità e gli eventi rari; tramite la componente rotata o ripetuta si stimano i parametri della popolazione perché incorpora i mutamenti avvenuti.

La varietà degli schemi che si possono adottare dipende anche dall'elaborazione teorica (Filippucci 1992) e dai limiti che l'organizzazione amministrativa pone alla loro realizzazione. La teoria investe sia la definizione dei termini, sia l'interrelazione tra economia e socio-demografia, sia la definizione dell'unità di rilevazione (la famiglia) che nella formulazione attuale pare inadeguata. In particolare, il concetto di capofamiglia si è modificato; le unità "co-residenziali" nella famiglia e i rapporti tra le famiglie determinano situazioni rilevanti nel consumo.

Una fase sperimentale sembra auspicabile per verificare sia l'entità degli errori non campionari, sia altri metodi di indagine che utilizzano il campionamento panel, rotato, *split panel*. Sarebbe altresì importante definire una strategia di rilevazione che si svincoli dagli apparati comunali e possa cogliere le modalità di acquisto nel tempo e nello spazio, il ciclo di vita, i rapporti tra le famiglie, l'incidenza delle variazioni dei prezzi di politica tariffaria, gli investimenti e i risparmi. In una indagine così complessa, tuttavia, la raccolta di un maggior numero di informazioni provoca un aumento delle mancate risposte. Una rilevazione accurata rimane, pertanto, difficile da realizzare, anche quando i concetti da misurare sono ben definiti e la struttura che segue il processo di raccolta dati è efficiente.

Riferimenti bibliografici

- CICCHITELLI G., HERZEL A., MONTANARI G. E. (1992), *Il campionamento statistico*, Bologna, il Mulino.
- COCHRAN W. G. (1977), *Sampling Techniques*, 3rd ed., New York, John Wiley & Sons.
- DIPPO C. S. e JACOBS C. A. (1983), *Area sample redesign for the consumer price index*, in "Proceedings of the Section on Survey Research Methods", American Statistical Association, 118-123.
- DRUDI I. (1992), *Analisi delle frequenze di acquisto nell'indagine sui bilanci di famiglia*, CON.PRI., Rapporto di ricerca n. 8, Bologna, Dipartimento di Scienze Statistiche.
- DUNCAN G. J. e KALTON G. (1987), *Issue of design and analysis of surveys across time*, in "International Statistic Review", 55, 97-117.
- EFRON B. (1994), *Missing data, imputation, and the bootstrap*, in "Journal of the American Statistical Association", 89, pp. 463-475.
- EFRON B. e TIBSHIRANI R. J. (1993), *An Introduction to the Bootstrap*, New York, Chapman & Hall.
- FALORSI P. D. e RUSSO A. (1992), *La mancata risposta totale nei campioni complessi: un'applicazione all'indagine campionaria sui consumi delle famiglie*, CON.PRI., Rapporto di ricerca n. 23, Bologna, Dipartimento di Scienze Statistiche.
- FALORSI P. D., FALORSI S., e RUSSO A. (1992), *Indagine campionaria sui consumi delle famiglie. Strategie di campionamento e precisione delle principali stime*, CON.PRI., Rapporto di ricerca n. 3, Bologna, Dipartimento di Scienze Statistiche.
- FILIPPUCCI C. (1992), *Sviluppi della teoria del consumo e problemi di misura*, in "Atti della XXXVI Riunione Scientifica della S.I.S.", vol. 2, pp. 11-18, Roma, Cisu.
- FILIPPUCCI C. e MARLIANI G. (1992a), *Un progetto di ricerca sulla misura dei consumi privati in Italia*, Progetto CON.PRI., nota interna n. 6, Roma, Istat.
- FILIPPUCCI C. e MARLIANI G. (1992b), *La misura del consumo delle famiglie: una riflessione a partire dall'esperienza italiana*, CON.PRI., Rapporto di ricerca n. 6, Bologna, Dipartimento di Scienze Statistiche.

Appendice A: tavole

Tavola A1 — Numero di famiglie intervistate ogni decade per regione e per mese nell'indagine sui consumi dell'Istat del 1989.

Regione	Gennaio			Febbraio			Marzo			Aprile		
	I	II	III	I	II	III	I	II	III	I	II	III
Piemonte	71	68	70	74	69	73	59	58	56	61	60	63
Valle d'Aosta	24	25	25	19	22	24	26	23	23	27	28	27
Lombardia	81	84	80	93	91	102	87	87	82	62	63	65
Trentino-Alto-Adige	30	30	30	30	29	30	111	116	112	30	30	30
Veneto	49	48	48	47	47	47	57	57	55	52	49	49
Friuli-Venezia-Giulia	25	37	29	25	25	25	30	30	30	30	31	31
Liguria	41	43	44	45	44	44	52	49	46	37	38	36
Emilia-Romagna	58	57	58	3	1	3	113	112	113	52	51	49
Toscana	63	66	66	66	67	67	64	61	60	64	68	66
Umbria	23	24	25	24	23	25	29	27	30	23	23	24
Marche	40	45	43	38	36	37	43	44	42	43	43	43
Lazio	72	70	74	76	77	73	75	73	77	74	72	70
Abruzzo	31	33	31	26	28	26	33	33	32	24	24	24
Molise	21	21	21	30	35	31	29	30	29	24	24	24
Campania	69	68	69	54	55	51	43	44	42	51	50	46
Puglia	48	47	47	48	48	48	58	59	59	57	57	53
Basilicata	30	30	30	29	30	30	25	23	23	27	28	26
Calabria	30	38	34	41	40	41	30	28	29	28	28	28
Sicilia	61	64	62	67	64	67	61	60	58	64	64	62
Sardegna	27	28	28	27	27	27	32	29	31	26	27	27
Totale	894	926	914	862	857	871	1080	1066	1055	856	858	847

(continua)

segue Tavola A1 — Numero di famiglie intervistate ogni decade per regione e per mese nell'indagine sui consumi dell'Istat del 1989.

Regione	Maggio			Giugno			Luglio			Agosto		
	I	II	III	I	II	III	I	II	III	I	II	III
Piemonte	74	75	72	56	55	56	58	62	62	79	69	78
Valle d'Aosta	24	24	24	24	23	22	22	20	21	29	28	27
Lombardia	105	109	107	103	99	100	75	71	72	95	101	95
Trentino-Alto-Adige	30	30	30	103	114	120	—	—	—	55	55	53
Veneto	48	45	46	36	35	36	50	48	48	50	49	48
Friuli-Venezia-Giulia	25	25	25	30	30	30	31	30	30	31	30	29
Liguria	48	48	45	45	42	41	45	44	44	33	32	35
Emilia-Romagna	60	61	61	51	54	51	53	54	54	62	58	59
Toscana	65	63	64	68	71	62	65	66	64	62	62	61
Umbria	30	30	30	26	27	27	28	26	28	23	23	22
Marche	35	36	39	43	43	43	42	44	43	44	42	43
Lazio	75	74	75	72	66	74	66	69	70	62	61	62
Abruzzo	29	29	29	29	27	31	22	23	24	32	32	32
Molise	32	32	32	24	24	24	27	29	28	32	31	29
Campania	72	78	75	53	54	54	64	62	65	52	51	52
Puglia	46	46	45	61	58	59	51	54	54	55	55	52
Basilicata	26	28	27	26	26	26	28	28	27	22	23	20
Calabria	30	32	32	26	23	36	34	37	37	30	31	29
Sicilia	71	69	71	55	52	55	52	49	49	68	65	65
Sardegna	35	35	35	31	31	31	29	29	29	30	30	28
Totale	960	969	964	962	954	977	842	845	849	946	928	909

(continua)

- FILIPPUCCI C. e FERRANTE M. R. (1993), *L'impiego dei diari individuali e dei documenti di spesa nell'indagine sui consumi*, CON.PRI., Rapporto di ricerca n. 11, Bologna, Dipartimento di Scienze Statistiche.
- FILIPPUCCI C. e ZANELLI E. (1994), *Un esperimento di valutazione cognitiva del questionario nell'indagine sui consumi*, CON.PRI., Rapporto di ricerca n. 16, Bologna, Dipartimento di Scienze Statistiche.
- HANEFELD U. (1984), *The German socio-economic panel*, in "Proceedings of the Social Statistics Section", American Statistical Association, pp. 117-124.
- INNOCENZI G. (1992), *Principali aspetti dell'indagine Istat sui consumi delle famiglie*, CON.PRI., Rapporto di ricerca n. 2, Dipartimento di Scienze Statistiche.
- KASPRZYK D., DUNCAN G. J., KALTON G., e SINGH M. P. (1989), *Panel Surveys*, New York, John Wiley & Sons.
- KISH L. (1965), *Sampling Survey*, New York, John Wiley & Sons.
- KISH L. (1983), *Data collection for details over space and time*, in Wright T., *Statistical Methods and the Improvement of Data Quality*, pp. 73-84, New York, Academic Press.
- KISH L. (1986), *Timing of surveys for public policy*, in "The Australian Journal of Statistics", 28, 1-12.
- KISH L. (1989), *Sampling Methods for Agricultural Surveys*, Roma, FAO Statistical Development Series, N. 3.
- LALLA M. (1992), *Metodi di indagine nella rilevazione dei consumi: proposte alternative all'indagine ripetuta*, CON.PRI., Rapporto di ricerca n. 5, Dipartimento di Scienze Statistiche.
- LIEPINS G. E. e UPPULURI V. R. R. (eds.) (1990), *Data Quality Control. Theory and Pragmatics*, New York, Marcel Dekker.
- LINDMAN H. R. (1991), *Analysis of Variance in Experimental Design*, New York, Springer-Verlag.
- LUCEV D. (1992a), *Le mancate risposte totali nell'indagine sui consumi delle famiglie*, CON.PRI., Rapporto di ricerca n. 14, Bologna, Dipartimento di Scienze Statistiche.
- LUCEV D. (1992b), *La revisione manuale dei dati nell'indagine sui consumi delle famiglie*, CON.PRI., Rapporto di ricerca n. 18, Bologna, Dipartimento di Scienze Statistiche.

- RICOLFI L. (1992), *HELGA: nuovi principi di analisi dei gruppi*, Milano, Franco Angeli.
- SÄRNDAL C. E. e DEVILLE G. C. (1992), *Calibration estimators in survey sampling*, in "Journal of the American Statistical Association", 87, pp. 376-382.
- SURVEY RESEARCH CENTER (1984), *User guide to the Panel Study of Income Dynamics*, Ann Arbor, Michigan, Inter-university Consortium for Political and Social Research.
- WOLTER K. M. e CAUSEY B. D. (1991), *Evaluation of procedures for improving population estimation*, in "Journal of the American Statistical Association", 86, pp. 278-284.
- ZIESCHANG K. D. (1990), *Sample weighting methods and estimation of totals in the consumer expenditure survey*, in "Journal of the American Statistical Association", 85, pp. 986-1001.

segue Tavola A1 — *Numero di famiglie intervistate ogni decade per regione e per mese nell'indagine sui consumi dell'Istat del 1989.*

Regione	Settembre			Ottobre			Novembre			Dicembre		
	I	II	III	I	II	III	I	II	III	I	II	III
Piemonte	50	50	51	62	64	62	79	69	65	63	55	61
Valle d'Aosta	31	27	26	29	28	25	24	24	22	31	25	31
Lombardia	93	94	105	81	76	78	76	83	78	117	123	125
Trentino-Alto-Adige	104	125	114	26	25	25	34	36	35	115	114	113
Veneto	46	54	55	50	50	48	48	48	46	54	48	55
Friuli-Venezia-Giulia	34	32	30	29	31	31	24	25	25	26	25	25
Liguria	47	50	48	46	49	42	46	42	50	47	45	54
Emilia-Romagna	57	55	57	54	52	54	57	56	57	60	60	60
Toscana	62	66	67	61	65	64	66	66	63	63	61	63
Umbria	38	33	34	19	20	20	27	23	28	38	36	38
Marche	44	41	45	41	45	40	43	41	43	43	43	43
Lazio	77	82	83	69	69	71	63	61	62	73	69	85
Abruzzo	29	28	28	31	28	30	29	27	36	32	31	32
Molise	23	23	23	27	26	28	30	30	32	20	20	18
Campania	57	56	57	64	62	61	56	61	59	59	57	56
Puglia	57	55	54	51	52	52	54	57	58	56	54	58
Basilicata	27	29	28	23	24	27	25	27	29	24	25	24
Calabria	28	27	39	32	31	31	37	37	36	38	37	51
Sicilia	43	42	45	55	52	54	66	63	66	58	55	59
Sardegna	29	28	31	29	30	28	29	27	31	30	29	30
Totale	975	987	1020	879	879	871	913	903	911	1047	1012	1080

(continua)

segue Tavola A1 — *Numero di famiglie intervistate per regione nell'indagine sui consumi dell'Istat del 1989: totale annuo, numero medio per decade (rispetto all'anno), e numero medio per tipo di decade.*

Regione				\bar{n}_w		
	Totale	\bar{n}_w	DS(.) ^(*)	I	II	III
Piemonte	2299	64	8	66	63	63
Valle d'Aosta	904	25	3	26	25	25
Lombardia	3237	90	16	89	90	91
Trentino-Alto-Adige	2064	63	40	61	64	63
Veneto	1745	48	5	49	48	48
Friuli-Venezia-Giulia	1031	29	3	28	29	28
Liguria	1587	44	5	44	44	44
Emilia-Romagna	2027	56	23	57	56	56
Toscana	2316	64	2	64	65	64
Umbria	974	27	5	27	26	28
Marche	1506	42	3	42	42	42
Lazio	2567	71	6	71	69	73
Abruzzo	1035	29	3	29	29	29
Molise	963	27	4	27	27	27
Campania	2151	60	8	60	60	59
Puglia	1923	53	4	54	54	53
Basilicata	950	26	3	26	27	26
Calabria	1196	33	6	32	32	35
Sicilia	2133	59	8	60	58	59
Sardegna	1060	29	2	30	29	30
Totale	33668	939	—	942	937	943

(*) DS(.) Deviazione Standard del numero di USS intervistate nelle diverse decadi dell'anno
 \bar{n}_w media del numero di USS intervistate nelle diverse decadi dell'anno

Tavola A2 — Numero di famiglie intervistate per regione, per altezza sul livello del mare delle UPS, e per attività economica prevalente nell'indagine sui consumi dell'Istat del 1989.

Regione	Altezza s. livello d. mare(*)			Attività economica prev. (*)		
	M.	C.	P.	Agric.	Ind.	Serv.
Piemonte	211	751	1301	130	703	1430
Valle d'Aosta	904	—	—	60	214	630
Lombardia	262	668	2250	—	1564	1616
Trentino-Alto-Adige	2064	—	—	—	289	1775
Veneto	177	244	1291	77	659	976
Friuli-Venezia-Giulia	141	434	456	—	217	814
Liguria	695	853	—	74	64	1410
Emilia-Romagna	54	475	1421	98	426	1426
Toscana	461	1405	433	—	661	1638
Umbria	156	800	—	—	356	600
Marche	108	1389	—	—	679	827
Lazio	289	763	1487	47	154	2338
Abruzzo	352	647	—	36	316	647
Molise	486	468	—	281	199	474
Campania	176	1461	454	340	117	1634
Puglia	—	485	1437	337	129	1456
Basilicata	482	349	76	297	242	368
Calabria	292	648	244	176	137	871
Sicilia	224	917	976	464	166	1487
Sardegna	—	526	522	34	143	870
Totale	7534	13292	12348	2452	7435	23287

(*) Ci sono 494 casi mancanti

Tavola A3 — Coefficiente di variazione della spesa per il consumo totale, per il consumo di beni alimentari, e per il consumo di beni non alimentari delle famiglie per regione e per trimestre nell'indagine sul consumo del 1989 (pesati con i coefficienti di riporto trimestrali).

Regione	Primo trimestre			Secondo trimestre			Terzo trimestre		
	Tot.	Al.	N.A.	Tot.	Al.	N.A.	Tot.	Al.	N.A.
Piemonte	86,6	57,7	102,0	83,4	52,6	99,5	81,2	59,2	96,1
Valle d'Aosta	79,3	65,8	91,6	128,9	59,9	158,0	94,4	51,8	115,8
Lombardia	75,5	59,0	89,0	91,1	60,9	107,8	98,5	58,5	116,3
Trentino-Alto-Adige	79,9	64,5	92,7	78,8	52,7	92,4	89,5	53,7	105,2
Veneto	75,8	64,4	876,6	88,2	66,6	100,2	73,3	51,8	85,4
Friuli-Venezia-Giulia	97,5	57,8	116,1	91,6	62,9	108,9	83,8	55,2	99,0
Liguria	70,0	50,0	85,9	70,6	52,8	85,1	92,2	64,0	108,4
Emilia-Romagna	85,1	70,9	99,3	86,6	64,7	101,5	79,6	60,6	91,6
Toscana	73,9	54,0	88,8	81,6	55,1	98,2	96,3	63,2	118,1
Umbria	73,2	57,4	88,3	80,7	70,0	93,1	87,1	51,5	106,6
Marche	79,5	56,2	95,0	90,9	54,5	112,2	72,6	54,7	86,1
Lazio	75,2	47,0	95,1	90,6	50,8	113,3	79,1	51,1	98,2
Abruzzo	81,5	59,9	96,7	100,2	65,2	117,9	74,9	50,8	89,5
Molise	97,6	55,1	122,8	77,9	53,7	95,9	77,4	60,6	95,0
Campania	77,3	51,6	100,8	70,9	51,1	89,4	83,1	51,7	113,7
Puglia	112,0	56,4	140,2	129,5	50,3	166,1	85,2	47,9	110,1
Basilicata	115,8	65,0	147,0	124,3	55,4	158,2	90,3	54,5	113,0
Calabria	96,1	58,1	122,6	161,0	56,4	213,5	78,9	58,5	100,5
Sicilia	88,8	60,2	108,9	84,2	54,5	106,1	92,1	51,0	118,2
Sardegna	85,3	59,6	104,5	90,1	52,1	113,5	87,5	50,2	107,6
Totale	84,0	58,4	101,8	94,3	57,9	114,8	89,8	56,2	109,5

(continua)

segue Tavola A3 — *Coefficiente di variazione della spesa per il consumo totale, per il consumo di beni alimentari, e per il consumo di beni non alimentari delle famiglie per regione e per trimestre nell'indagine sul consumo del 1989 (pesati con i coefficienti di riporto trimestrali).*

Regione	Quarto trimestre			Totale "campione-anno"		
	Tot.	Alim.	N.A.	Tot.	Alim.	N.A.
Piemonte	74,8	51,5	89,0	81,5	55,5	96,6
Valle d'Aosta	84,5	57,9	102,3	98,4	59,0	119,7
Lombardia	81,9	61,3	95,6	88,9	60,3	104,8
Trentino-Alto-Adige	85,8	58,2	102,6	85,1	57,5	99,3
Veneto	80,5	51,5	98,3	81,6	58,7	94,3
Friuli-Venezia-Giulia	75,8	54,5	98,0	89,2	57,9	105,5
Liguria	81,8	55,6	98,9	81,6	56,1	97,5
Emilia-Romagna	79,8	57,0	91,6	83,0	63,5	96,0
Toscana	88,1	56,0	106,2	86,0	57,2	103,4
Umbria	69,2	66,5	79,2	78,4	62,9	93,1
Marche	74,2	57,4	89,1	79,5	56,2	95,9
Lazio	71,8	51,9	85,9	79,7	50,8	98,6
Abruzzo	99,6	60,0	117,0	91,8	59,6	108,6
Molise	61,9	56,1	72,9	80,8	58,6	100,1
Campania	67,7	55,1	84,2	75,0	52,5	97,4
Puglia	95,7	57,8	120,4	107,7	53,6	137,9
Basilicata	83,7	62,0	100,4	104,0	60,9	130,2
Calabria	91,7	68,3	112,5	107,8	62,9	139,6
Sicilia	89,9	52,3	113,4	89,2	54,9	112,1
Sardegna	93,7	63,7	110,9	95,5	59,2	116,2
Totale	85,4	57,6	102,2	88,8	57,8	107,5

Tavola A4 — *Numero di famiglie, media e deviazione standard delle spese per il consumo totale, per regione e per trimestre nell'indagine Istat del 1989 (pesati con i coefficienti di riporto trimestrali).*

Regione		Primo trimestre			Secondo trimestre		
		n	\bar{Y}	D.S.	n	\bar{Y}	D.S.
Piemonte		278	1877,2	1338,5	252	2163,8	1790,5
	NAR	320	2444,3	2180,0	320	2356,7	1966,0
Valle d'Aosta		67	2345,5	1914,0	80	3138,5	4712,7
	NAR	144	2506,9	1589,5	143	1810,3	1160,5
Lombardia		423	2289,7	1774,8	455	2723,6	2861,0
	NAR	364	2418,5	1799,5	358	2826,2	2366,3
Trentino-Alto-Adige		126	2149,7	1653,0	124	2238,3	1505,1
	NAR	392	2565,1	2050,5	393	2680,5	2162,1
Veneto		193	2139,6	1766,7	174	2081,1	1601,6
	NAR	262	2553,7	1864,9	221	2876,7	2536,5
Friuli-Venezia-Giulia		117	2471,9	2818,2	116	2400,4	2324,9
	NAR	139	2129,8	1715,2	141	2191,6	1910,6
Liguria		275	1928,2	1396,5	265	2028,5	1315,6
	NAR	133	2031,7	1358,0	115	2119,7	1625,6
Emilia-Romagna		321	2558,5	2081,4	312	2709,9	2416,3
	NAR	197	2631,8	2315,7	178	2575,1	2162,9
Toscana		326	2374,6	1597,6	323	2313,6	1892,0
	NAR	253	2276,9	1793,9	268	2355,9	1920,8
Umbria		92	2434,1	1944,5	115	2022,6	1898,8
	NAR	138	1831,7	1099,2	125	1826,8	1252,8
Marche		162	3104,1	2271,5	162	3051,1	2355,3
	NAR	206	2425,9	1959,7	206	2308,6	2203,6
Lazio		509	2206,9	1762,9	515	2372,3	2019,5
	NAR	158	2161,9	1434,5	141	2246,2	2234,6
Abruzzo		108	2210,8	1278,0	81	2532,2	2634,1
	NAR	165	2096,5	1854,4	165	2164,1	2114,0
Molise		64	2833,0	3008,7	63	2486,9	1111,2
	NAR	183	1636,0	1273,8	177	1520,1	1317,9
Campania		342	1730,8	1383,2	310	1716,8	1217,1
	NAR	225	1852,5	1397,9	223	1724,7	1222,9
Puglia		197	1904,4	1303,7	235	1911,8	1731,6
	NAR	265	2020,0	2553,3	247	2091,6	2977,4

(continua)

segue Tavola A4 — Numero di famiglie, media e deviazione standard delle spese per il consumo totale, per regione e per trimestre nell'indagine Istat del 1989 (pesati con i coefficienti di riporto trimestrali).

Regione		Primo trimestre			Secondo trimestre		
		n	\bar{Y}	D.S.	n	\bar{Y}	D.S.
Basilicata		46	1625,0	1785,3	36	1571,9	1173,3
	NAR	204	1618,5	1893,1	204	1682,0	2206,9
Calabria		135	1509,5	1306,4	116	1574,1	910,8
	NAR	176	1718,1	1682,4	147	1633,3	2943,4
Sicilia		312	2334,5	1891,8	311	2110,4	1267,8
	NAR	252	1587,4	1422,2	252	1558,7	1542,3
Sardegna		108	2377,4	1562,2	108	2850,9	2195,2
	NAR	148	1513,1	1375,2	170	1717,7	1550,4
Totale		4201	2175,8	1747,1	4153	2284,8	2039,8
	NAR	4324	2168,0	1868,1	4194	2262,3	2200,0
Totale: AR + NAR		8525	2171,0	1823,3	8347	2270,8	2141,0

(continua)

segue Tavola A4 — Numero di famiglie, media e deviazione standard delle spese per il consumo totale, per regione e per trimestre nell'indagine Istat del 1989 (pesati con i coefficienti di riporto trimestrali).

Regione		Terzo trimestre			Quarto trimestre		
		n	\bar{Y}	D.S.	n	\bar{Y}	D.S.
Piemonte		271	2157,4	1630,4	278	2482,7	2506,2
	NAR	278	2482,7	2506,2	305	2343,7	1677,5
Valle d'Aosta		81	2617,3	1828,8	80	2486,4	2061,6
	NAR	150	2419,8	2537,8	159	2448,0	2086,5
Lombardia		444	2627,7	2401,7	482	2829,3	2419,6
	NAR	356	3028,8	3040,3	355	3164,5	2554,7
Trentino-Alto-Adige		128	2326,0	1390,8	126	2610,3	2019,7
	NAR	378	2777,7	2627,1	397	2956,7	2691,5
Veneto		177	2162,5	1566,1	188	2562,1	1768,0
	NAR	271	2781,3	2002,4	259	3109,6	2709,4
Friuli-Venezia-Giulia		117	2177,8	2120,8	117	2368,3	1689,0
	NAR	160	2625,9	2002,6	124	3041,1	2608,7

(continua)

segue Tavola A4 — Numero di famiglie, media e deviazione standard delle spese per il consumo totale, per regione e per trimestre nell'indagine Istat del 1989 (pesati con i coefficienti di riporto trimestrali).

Regione		Terzo trimestre			Quarto trimestre		
		n	\bar{Y}	D.S.	n	\bar{Y}	D.S.
Liguria		230	2037,4	1743,2	259	2316,1	2002,4
	NAR	148	2572,1	2442,5	162	2643,5	2119,5
Emilia-Romagna		298	2710,3	2126,3	316	2897,7	2224,4
	NAR	211	2925,9	2348,8	194	3251,1	2642,9
Toscana		351	2742,7	2924,3	351	2713,3	2436,8
	NAR	224	2183,6	1722,0	220	2558,5	2210,5
Umbria		97	2246,7	2221,3	102	2866,0	2055,4
	NAR	158	2157,2	1685,3	147	2456,0	1617,6
Marche		163	2916,3	1692,5	162	2797,8	1745,1
	NAR	225	2376,1	1842,1	220	2695,9	2097,6
Lazio		482	2355,9	1686,7	475	3048,5	2118,1
	NAR	140	2256,6	2048,9	147	2230,0	1555,5
Abruzzo		75	2572,7	1913,8	104	2632,9	1927,0
	NAR	175	2371,9	1777,9	162	2843,9	3006,4
Molise		64	2929,3	1505,9	62	2653,2	1205,6
	NAR	181	2258,0	1907,6	169	1763,7	1134,8
Campania		275	1850,5	1903,8	296	1847,2	1226,9
	NAR	241	1706,6	1056,3	239	1844,4	1265,9
Puglia		235	1975,2	1275,6	222	2244,5	1687,8
	NAR	252	1851,8	1754,7	270	1862,3	1989,2
Basilicata		49	1252,4	722,9	45	2089,1	1862,6
	NAR	183	1648,8	1524,0	183	1985,4	1634,3
Calabria		110	1902,6	1297,1	132	1964,8	1652,1
	NAR	182	1967,1	1609,2	198	1986,8	1866,7
Sicilia		263	1985,0	1495,0	287	2583,7	2266,9
	NAR	215	1700,1	1732,7	241	1682,3	1384,0
Sardegna		97	3246,5	2522,6	98	3548,0	2428,2
	NAR	166	2452,9	2224,5	165	2953,6	3074,5
Totale		4007	2323,1	1998,9	4209	2604,2	2076,6
	NAR	4294	2389,0	2195,7	4286	2553,6	2267,5
Totale: AR + NAR		8301	2364,1	2123,9	8495	2572,6	2197,8

Tavola A5 — Numero di famiglie, media e deviazione standard delle spese per il consumo di beni alimentari, per regione e per trimestre nell'indagine Istat del 1989 (pesati con i coefficienti di riporto trimestrali).

Regione		Primo trimestre			Secondo trimestre		
		<i>n</i>	\bar{Y}	D.S.	<i>n</i>	\bar{Y}	D.S.
Piemonte		277	446,3	271,2	247	512,0	271,4
	NAR	317	476,7	267,2	318	503,0	263,4
Valle d'Aosta		62	578,9	358,6	75	596,4	330,3
	NAR	143	476,9	318,2	142	450,5	268,6
Lombardia		415	531,7	316,6	441	554,2	327,9
	NAR	363	558,2	327,5	354	595,6	365,9
Trentino-Alto-Adige		124	458,8	316,9	123	447,1	238,6
	NAR	391	459,5	289,0	390	477,9	250,2
Veneto		191	472,2	276,4	172	457,4	252,1
	NAR	258	525,2	345,8	220	556,5	380,1
Friuli-Venezia-Giulia		117	456,1	257,3	116	505,1	310,2
	NAR	136	460,3	269,6	140	486,5	310,0
Liguria		274	518,4	253,2	263	516,9	272,4
	NAR	132	458,9	233,3	115	502,1	265,8
Emilia-Romagna		317	542,7	407,9	305	537,8	375,8
	NAR	194	486,4	317,8	177	535,3	320,4
Toscana		323	588,6	339,9	320	616,3	355,4
	NAR	252	538,9	268,4	268	560,3	292,0
Umbria		92	582,5	365,4	115	557,5	403,1
	NAR	138	549,7	292,8	125	528,9	360,8
Marche		162	703,8	349,0	161	703,5	318,4
	NAR	204	580,7	336,1	203	562,0	320,1
Lazio		504	593,2	289,9	513	579,7	278,1
	NAR	158	581,5	254,0	141	514,4	283,7
Abruzzo		108	546,5	261,0	81	532,5	328,3
	NAR	165	503,6	321,0	165	489,4	324,3
Molise		64	598,8	212,6	63	594,9	213,4
	NAR	181	436,2	259,4	176	419,9	239,3
Campania		341	544,0	309,0	310	556,4	312,8
	NAR	225	600,3	287,5	221	534,0	251,2
Puglia		196	503,6	210,9	235	538,4	225,6
	NAR	265	485,4	304,2	247	481,3	261,3

(continua)

segue Tavola A5 — Numero di famiglie, media e deviazione standard delle spese per il consumo di beni alimentari, per regione e per trimestre nell'indagine Istat del 1989 (pesati con i coefficienti di riporto trimestrali).

Regione		Primo trimestre			Secondo trimestre		
		<i>n</i>	\bar{Y}	D.S.	<i>n</i>	\bar{Y}	D.S.
Basilicata		46	352,6	128,2	35	385,6	158,6
	NAR	203	449,6	301,2	202	439,3	250,6
Calabria		135	361,7	229,7	114	386,8	271,7
	NAR	175	501,6	275,5	147	458,2	239,2
Sicilia		308	588,6	331,0	305	586,1	283,6
	NAR	250	425,4	248,1	251	411,5	218,9
Sardegna		108	497,7	273,0	108	576,9	298,5
	NAR	146	477,0	292,0	170	484,3	248,0
Totale		4164	535,4	312,7	4102	548,3	300,6
	NAR	4296	515,5	300,8	4172	519,5	305,5
Totale: AR + NAR		8460	523,0	305,5	8274	530,3	307,0

segue Tavola A5 — Numero di famiglie, media e deviazione standard delle spese per il consumo di beni alimentari, per regione e per trimestre nell'indagine Istat del 1989 (pesati con i coefficienti di riporto trimestrali).

Regione		Terzo trimestre			Quarto trimestre		
		<i>n</i>	\bar{Y}	D.S.	<i>n</i>	\bar{Y}	D.S.
Piemonte		264	467,2	241,6	303	507,7	247,3
	NAR	276	514,1	318,6	275	575,1	299,6
Valle d'Aosta		79	572,6	298,3	76	570,3	320,2
	NAR	147	552,1	284,7	158	558,8	327,4
Lombardia		428	539,4	342,5	477	563,9	374,2
	NAR	346	604,5	338,7	352	655,6	384,0
Trentino-Alto-Adige		127	461,8	268,4	126	492,1	253,0
	NAR	376	487,4	254,4	396	502,0	302,4
Veneto		171	465,7	247,8	188	542,6	331,6
	NAR	269	575,7	289,1	256	556,1	263,2
Friuli-Venezia-Giulia		117	492,0	274,2	116	485,8	239,4
	NAR	160	472,6	259,2	123	534,0	302,0

(continua)

segue **Tavola A5** — Numero di famiglie, media e deviazione standard delle spese per il consumo di beni alimentari, per regione e per trimestre nell'indagine Istat del 1989 (pesati con i coefficienti di riporto trimestrali).

Regione		Terzo trimestre			Quarto trimestre		
		n	\bar{Y}	D.S.	n	\bar{Y}	D.S.
Liguria		229	487,4	293,1	257	511,4	289,2
	NAR	148	560,2	373,3	162	542,1	294,0
Emilia-Romagna		294	499,0	307,4	308	548,2	293,5
	NAR	208	552,9	328,2	193	660,2	379,0
Toscana		346	597,4	364,6	351	650,5	362,5
	NAR	220	556,0	359,7	218	541,8	294,1
Umbria		97	573,4	335,1	102	811,6	522,9
	NAR	158	508,4	225,8	147	538,9	313,6
Marche		161	682,9	303,3	162	763,1	354,2
	NAR	217	579,5	335,6	218	653,2	398,1
Lazio		481	579,4	285,3	472	664,3	337,9
	NAR	139	550,5	298,2	146	630,1	336,3
Abruzzo		74	598,0	309,8	103	640,0	324,9
	NAR	174	590,5	298,0	162	618,2	388,3
Molise		64	713,8	250,1	62	708,8	254,0
	NAR	181	566,7	378,9	169	542,9	330,8
Campania		272	534,7	292,2	296	575,1	330,5
	NAR	240	601,3	297,0	238	571,6	305,2
Puglia		235	573,8	232,0	222	608,6	286,4
	NAR	252	517,7	266,0	270	535,3	337,4
Basilicata		49	388,9	183,1	45	621,5	419,0
	NAR	181	428,0	237,8	182	511,0	303,2
Calabria		110	592,5	438,1	132	547,9	490,9
	NAR	182	593,6	313,3	196	576,3	349,8
Sicilia		260	573,0	282,0	287	645,4	328,0
	NAR	214	449,4	221,0	240	497,7	248,3
Sardegna		96	693,7	268,4	98	686,3	455,9
	NAR	165	611,5	332,8	165	657,6	410,5
Totale		3954	543,3	304,9	4183	593,0	342,8
	NAR	4253	553,0	311,3	4266	585,3	336,0
Totale: AR + NAR		8207	549,3	309,0	8449	588,2	338,6

Tavola A6 — Numero di famiglie, media e deviazione standard delle spese per il consumo di beni non alimentari, per regione e per trimestre nell'indagine Istat del 1989 (pesati con i coefficienti di riporto trimestrali).

Regione		Primo trimestre			Secondo trimestre		
		n	\bar{Y}	D.S.	n	\bar{Y}	D.S.
Piemonte		278	1432,2	1199,9	252	1663,5	1680,9
	NAR	320	1973,6	2052,5	320	1858,9	1830,0
Valle d'Aosta		67	1809,6	1649,0	80	2598,0	4604,2
	NAR	144	1583,4	1443,8	143	1366,6	1026,4
Lombardia		423	1769,6	1642,7	455	2185,9	2744,5
	NAR	364	1863,2	1622,2	358	2238,4	2195,0
Trentino-Alto-Adige		126	1697,6	1468,3	124	1794,2	1404,9
	NAR	392	2106,7	1960,5	393	2209,3	2086,6
Veneto		193	1676,0	1656,4	174	1634,1	1453,2
	NAR	262	2040,5	1677,2	221	2323,3	2317,4
Friuli-Venezia-Giulia		117	2015,8	2705,0	116	1895,3	2165,6
	NAR	139	1681,4	1612,8	141	1708,4	1781,1
Liguria		275	1411,8	1265,1	265	1517,3	1182,2
	NAR	133	1583,3	1285,2	115	1617,6	1494,8
Emilia-Romagna		321	2025,8	1898,2	312	2186,3	2254,6
	NAR	197	2152,1	2221,5	178	2044,5	2036,4
Toscana		326	1790,1	1413,9	323	1700,4	1679,2
	NAR	253	1742,6	1670,4	268	1795,6	1752,9
Umbria		92	1851,7	1749,0	115	1465,1	1606,3
	NAR	138	1282,0	907,6	125	1297,9	997,5
Marche		162	2400,4	2105,0	162	2353,5	2247,5
	NAR	206	1851,7	1789,0	206	1756,1	2074,1
Lazio		509	1617,3	1653,5	515	1793,7	1945,1
	NAR	158	1580,4	1285,2	141	1731,8	2101,4
Abruzzo		108	1664,4	1110,2	81	1999,7	2485,9
	NAR	165	1592,9	1678,1	165	1674,7	1905,9
Molise		64	2234,3	2995,7	63	1892,0	1027,4
	NAR	183	1208,4	1139,0	177	1105,2	1204,3
Campania		342	1187,9	1251,8	310	1160,5	1051,5
	NAR	225	1252,2	1223,5	223	1191,6	1054,6
Puglia		197	1406,8	1219,7	235	1373,4	1673,6
	NAR	265	1534,5	2409,5	247	1610,3	2876,5

(continua)

segue Tavola A6 — Numero di famiglie, media e deviazione standard delle spese per il consumo di beni non alimentari, per regione e per trimestre nell'indagine Istat del 1989 (pesati con i coefficienti di riporto trimestrali).

Regione	Primo trimestre			Secondo trimestre		
	n	\bar{Y}	D.S.	n	\bar{Y}	D.S.
Basilicata	46	1272,4	1736,0	36	1195,5	1039,5
NAR	204	1171,3	1749,6	204	1249,2	2102,1
Calabria	135	1147,8	1186,2	116	1194,3	753,9
NAR	176	1222,6	1557,4	147	1175,1	2856,4
Sicilia	312	1751,1	1703,3	311	1537,6	1081,8
NAR	252	1164,9	1321,7	252	1149,4	1476,6
Sardegna	108	1879,6	1449,5	108	2273,9	2037,9
NAR	148	1051,5	1192,0	170	1233,5	1471,3
Totale	4201	1645,6	1606,5	4153	1744,1	1913,4
NAR	4324	1657,4	1727,4	4194	1746,1	2056,7
Totale: AR + NAR	8525	1652,9	1682,8	8347	1745,3	2003,8

(continua)

segue Tavola A6 — Numero di famiglie, media e deviazione standard delle spese per il consumo di beni non alimentari, per regione e per trimestre nell'indagine Istat del 1989 (pesati con i coefficienti di riporto trimestrali).

Regione	Terzo trimestre			Quarto trimestre		
	n	\bar{Y}	D.S.	n	\bar{Y}	D.S.
Piemonte	271	1702,5	1532,3	305	1837,3	1553,4
NAR	278	1973,1	1931,0	275	1942,5	1766,7
Valle d'Aosta	81	2057,7	1689,1	80	1951,7	1917,8
NAR	150	1878,0	2444,1	159	1892,8	1970,4
Lombardia	444	2105,1	2272,3	482	2270,3	2256,8
NAR	356	2439,6	2894,6	355	2514,8	2358,4
Trentino-Alto-Adige	128	1873,3	1264,0	126	2118,2	1926,6
NAR	378	2294,3	2552,7	397	2455,9	2576,0
Veneto	177	1711,5	1433,9	188	2019,5	1616,5
NAR	271	2211,4	1864,8	259	2567,5	2592,7
Friuli-Venezia-Giulia	117	1685,9	1984,2	117	1888,5	1563,7
NAR	160	2153,3	1922,1	124	2511,3	2520,1

(continua)

segue Tavola A6 — Numero di famiglie, media e deviazione standard delle spese per il consumo di beni non alimentari, per regione e per trimestre nell'indagine Istat del 1989 (pesati con i coefficienti di riporto trimestrali).

Regione	Terzo trimestre			Quarto trimestre		
	n	\bar{Y}	D.S.	n	\bar{Y}	D.S.
Liguria	230	1552,2	1614,3	259	1811,5	1865,5
NAR	148	2011,9	2187,1	162	2101,3	1968,5
Emilia-Romagna	298	2222,2	1997,1	316	2367,5	2093,7
NAR	211	2382,8	2208,0	194	2598,6	2428,6
Toscana	351	2154,4	2821,6	351	2062,8	2280,7
NAR	224	1635,6	1527,0	220	2020,8	2074,1
Umbria	97	1673,3	2042,6	102	2054,4	1687,0
NAR	158	1648,8	1575,5	147	1917,0	1473,2
Marche	163	2243,5	1536,6	162	2034,7	1581,5
NAR	225	1819,2	1679,7	220	2049,3	1893,7
Lazio	482	1777,3	1590,3	475	2388,2	1949,9
NAR	140	1709,8	1909,2	147	1603,0	1371,6
Abruzzo	75	1981,2	1694,2	104	2005,8	1766,9
NAR	175	1782,9	1618,6	162	2225,7	2742,6
Molise	64	2215,4	1449,8	62	1944,4	1023,0
NAR	181	1691,3	1751,8	169	1220,8	934,8
Campania	275	1319,3	1832,8	296	1272,1	1075,2
NAR	241	1106,9	875,8	238	1276,4	1072,6
Puglia	235	1401,4	1198,1	222	1635,9	1616,1
NAR	252	1334,2	1618,8	270	1327,0	1761,9
Basilicata	49	863,5	626,1	45	1467,6	1555,0
NAR	183	1227,1	1409,3	183	1476,7	1465,0
Calabria	110	1310,1	996,4	132	1416,9	1392,4
NAR	182	1373,5	1461,6	198	1416,1	1651,3
Sicilia	263	1419,3	1390,3	287	1938,4	2097,1
NAR	215	1252,9	1626,8	241	1186,9	1262,7
Sardegna	97	2562,4	2439,8	98	2861,8	2219,9
NAR	166	1846,0	2065,9	165	2296,0	2873,7
Totale	4007	1787,6	1887,2	4209	2015,2	1922,3
NAR	4294	1842,0	2057,4	4286	1972,5	2095,7
Totale: AR + NAR	8301	1821,5	1995,1	8495	1988,5	2032,3

Appendice B: grafici

Grafico 1 – Coefficiente di variazione della spesa per il consumo totale in ogni UPS rispetto alla loro dimensione (in centinaia di migliaia).

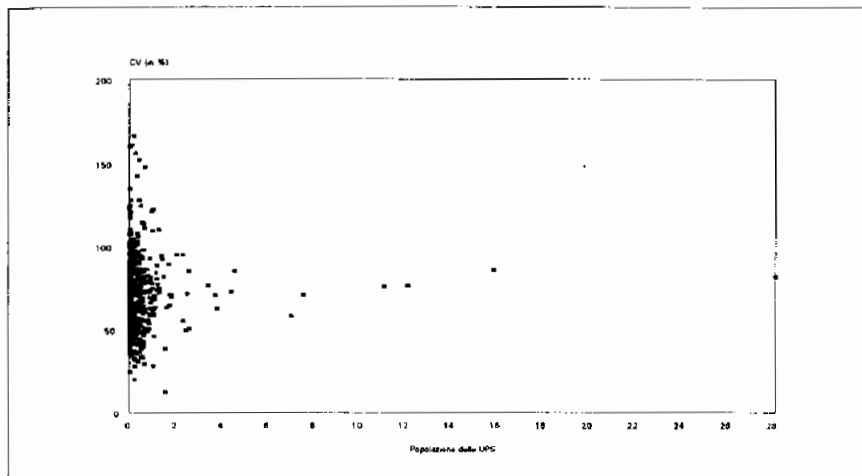


Grafico 2 – Coefficiente di variazione della spesa per il consumo totale in ogni UPS rispetto alla loro dimensione (in decine di migliaia).

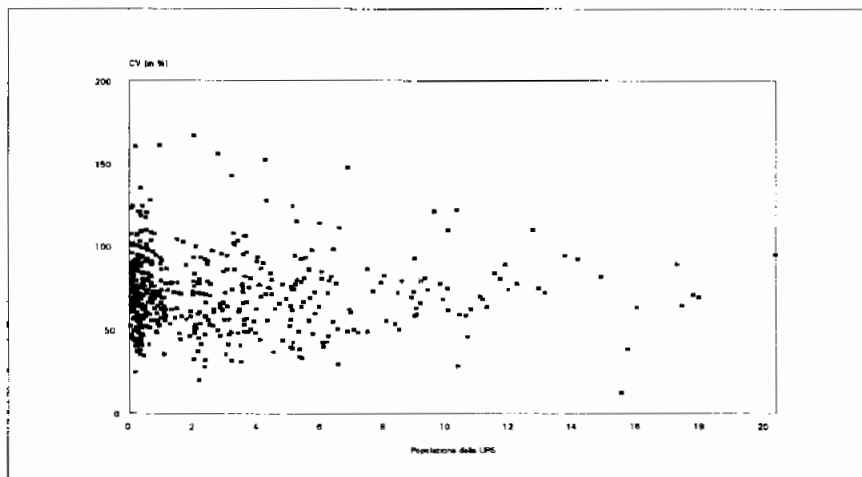


Grafico 3 – Coefficiente di variazione della spesa per il consumo totale di beni alimentari in ogni UPS rispetto alla loro dimensione (in centinaia di migliaia).

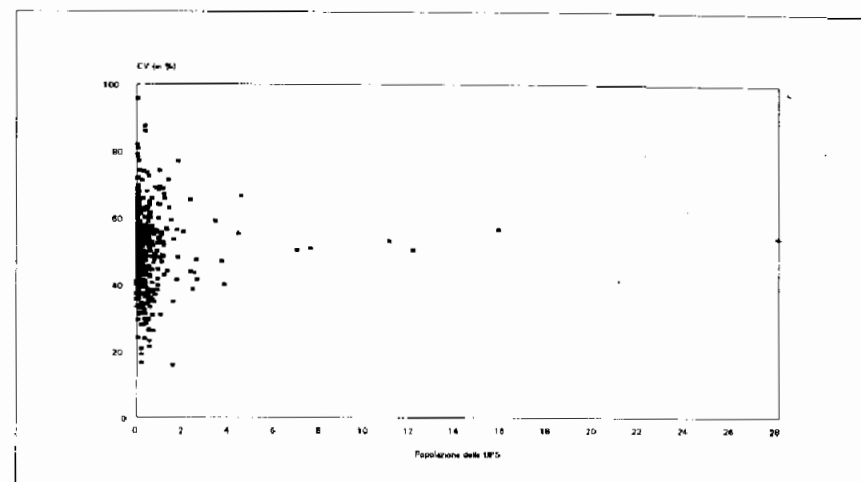


Grafico 4 – Coefficiente di variazione della spesa per il consumo totale di beni alimentari in ogni UPS rispetto alla loro dimensione (in decine di migliaia).

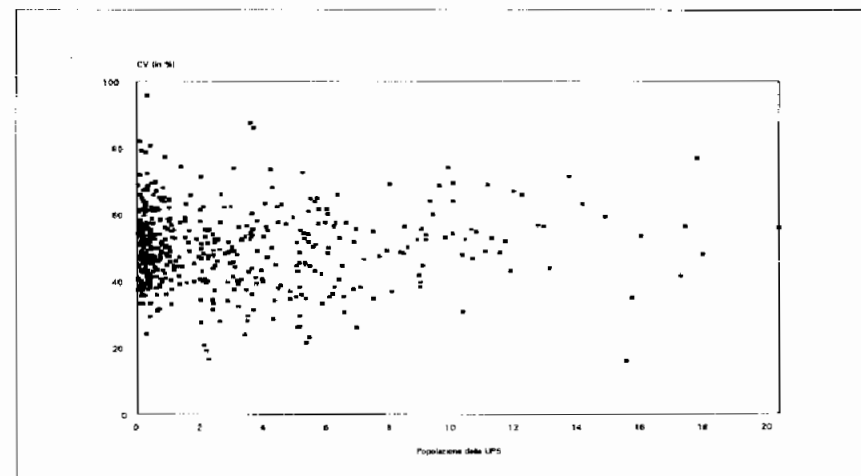


Grafico 5 – Coefficiente di variazione della spesa per il consumo totale di beni non alimentari in ogni UPS rispetto alla loro dimensione (in centinaia di migliaia)

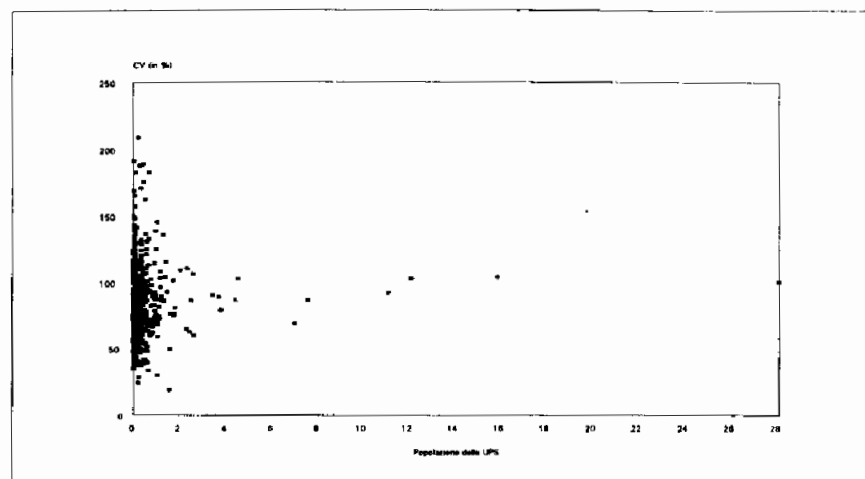


Grafico 6 – Coefficiente di variazione della spesa per il consumo totale di beni non alimentari in ogni UPS rispetto alla loro dimensione delle UPS (in decine di migliaia)

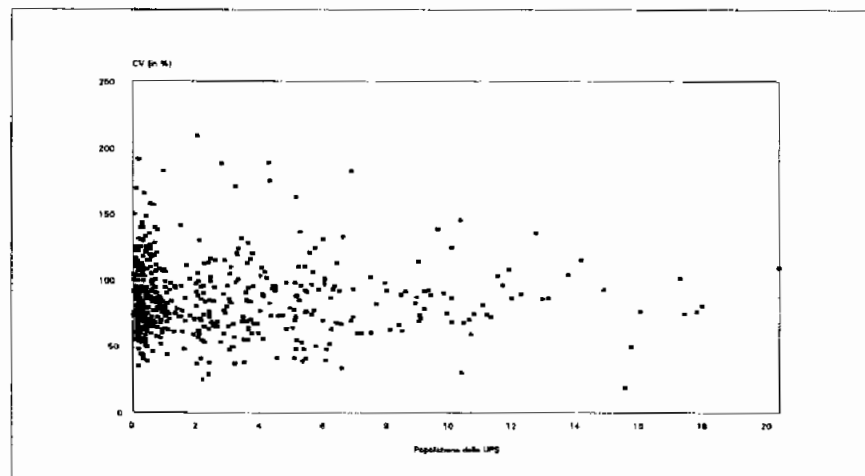
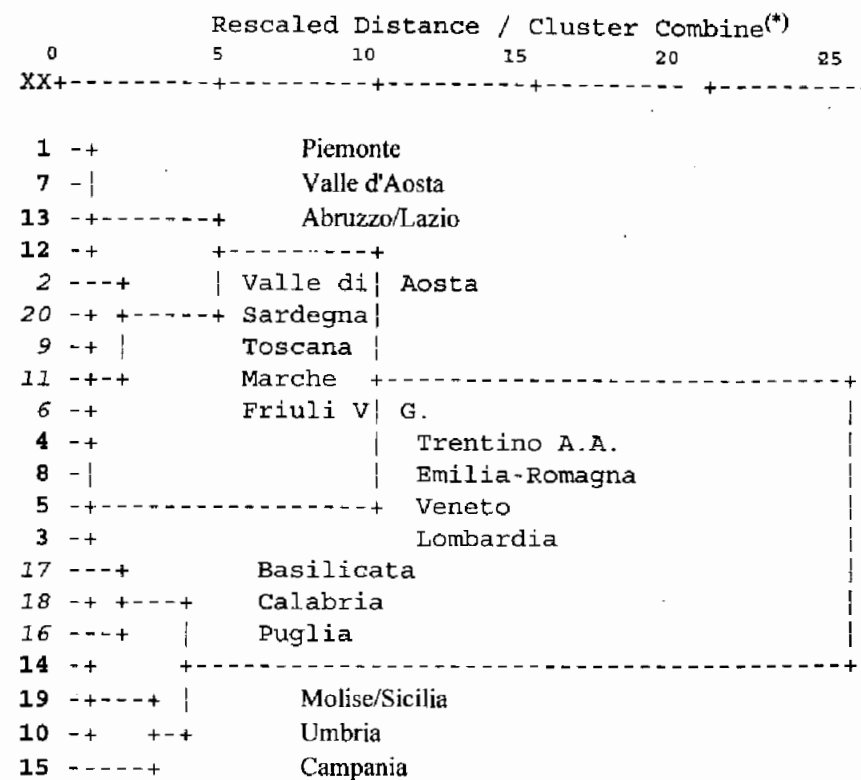
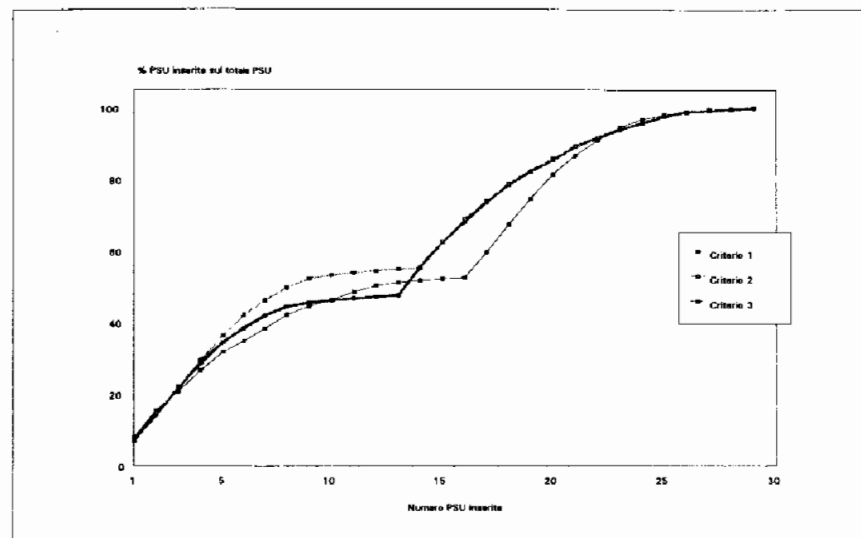


Grafico 7– Diagramma di formazione dei gruppi (dendrogramma) delle regioni ottenuto con il metodo di Ward in base alla media e alla deviazione standard della spesa per il consumo totale, e alla differenza delle spese per il consumo totale non alimentare e alimentare.



(*) I codici della prima "colonna" corrispondono ai codici delle regioni; il corpo e il tipo differente dei caratteri evidenziano i diversi gruppi risultanti dall'analisi dei gruppi: il primo gruppo è costituito da Piemonte, Liguria, Abruzzo, Lazio; il secondo da Valle d'Aosta, Sardegna, Toscana, Marche, Friuli-Venezia-Giulia; il terzo da Trentino-Alto-Adige, Emilia-Romagna, Veneto, Lombardia; il quarto da Basilicata, Calabria, Puglia; e il quinto da Molise, Sicilia, Umbria, Campania.

Grafico 8 – Percentuale delle UPS nel campione al variare del numero di UPS selezionate successivamente in base ai tre criteri di inclusione



Le legende "Criterio 1", "Criterio 2", e "Criterio 3" del grafico sopra riportato, e dei successivi, si riferiscono ai criteri di inclusione delle UPS utilizzati nei processi di simulazione (si veda il paragrafo 5.2).

Criterio 1: le UPS, ordinate in senso decrescente rispetto al numero di abitanti, vengono incluse successivamente fino all'esaurimento delle unità AR nella regione che ne ha di più; quindi, si procede con l'inserimento delle unità NAR seguendo la stessa regola (si veda il paragrafo 5.2.1).

Criterio 2: la prima UPS inclusa in ogni regione è la più popolata; le UPS rimanenti, ordinate in senso crescente rispetto al numero di abitanti, vengono incluse successivamente fino all'esaurimento delle unità NAR nella regione che ne ha di più; quindi, si procede con l'inserimento delle unità AR seguendo la stessa regola (si veda il paragrafo 5.2.2).

Criterio 3: le UPS, ordinate in senso crescente rispetto al numero di abitanti, vengono incluse successivamente fino all'esaurimento delle unità NAR nella regione che ne ha di più; quindi, si procede con l'inserimento delle unità AR seguendo la stessa regola (si veda il paragrafo 5.2.3).

Grafico 9 – Stime della spesa per il consumo totale al variare del numero di UPS nel campione in base ai tre criteri di inclusione

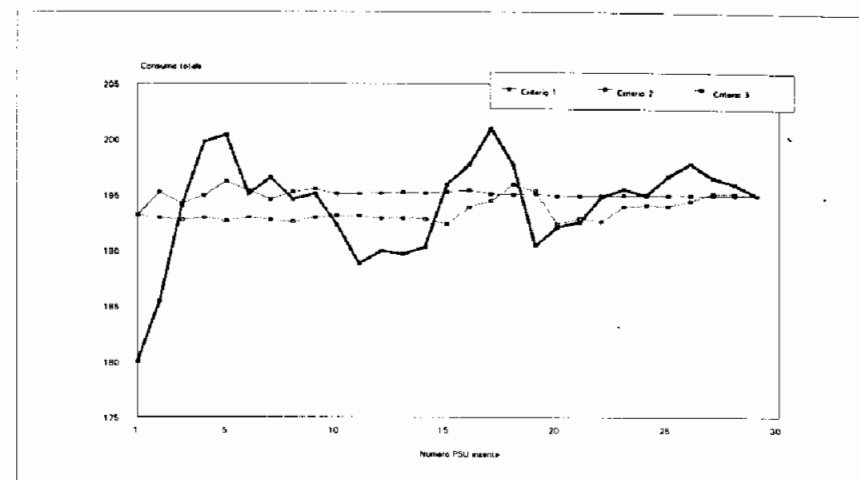


Grafico 10 – Deviazioni standard delle stime della spesa per il consumo totale al variare del numero di UPS inserite nel campione in base ai tre criteri di inclusione

